DEVELOPMENT OF NOVEL METHODS FOR IMAGING MEMBRANE PROTEIN STOICHIOMETRY AT THE SINGLE MOLECULE LEVEL AND FOR FABRICATION OF A SINGLE MOLECULE PROTEIN-DNA IMAGING DEVICE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Alexander LeRoy Van Slyke

August 2019

DEVELOPMENT OF NOVEL METHODS FOR IMAGING MEMBRAME PROTEIN STOICHIOMETRY AT THE SINGLE MOLECULE LEVEL AND FOR FABRICATION OF A SINGLE MOLECULE PROTEIN-DNA IMAGING DEVICE

Alexander LeRoy Van Slyke, Ph.D.
Cornell University 2019

Single molecule fluorescence imaging techniques have revolutionized the way we study biology by offering methods to examine the behavior and arrangement of individual molecules and the molecular mechanisms underlying biological processes. These techniques permit the investigation of transient states and critical heterogeneities undetectable by ensemble measurements and genome-wide assays. In this dissertation, I detail a series of projects, performed as a member of the laboratory of Warren Zipfel, in which new single molecule methods were created and applied to address important biological questions.

A technique designed to study protein-DNA interactions at the single molecule level in a high-throughput fashion is called “DNA curtains.” The nanopatterned microfluidic devices necessary for these experiments have previously been fabricated using electron-beam lithography. We developed a simplified, cost effective, and more accessible method of fabricating these devices.

Due to their modular DNA binding domain, transcription activator-like effectors (TALEs) have potential to be used in the study of gene function and

gene editing with medical and agricultural applications. Understanding the target search mechanism of TALEs is important to developing more efficient and accurate ways to design and deliver TALE proteins. In my first project, we investigated TALEs using “DNA curtains” in an effort to elucidate the details of this search mechanism.

Many single molecule techniques require the sample to be observed *in vitro* in order to isolate the biomolecule of interest. As a result, physiological behavior may not be preserved. In my second project, we developed a method named Single Protein Recovery After Dilution (SPReAD) that addresses this limitation by enabling protein stoichiometry and function to be studied *in vivo.*

My final project is an investigation of the functional composition of metabotropic glutamate receptors (mGluRs). Glutamate acts as both a neurotransmitter and neuromodulator in the central nervous system. These neuromodulatory effects are mediated by mGluRs and their improper function has been linked to schizophrenia and Fragile X Syndrome. Understanding the stoichiometry of mGluR complexes is necessary to the development of pharmacological compounds which modulate their signaling. We investigated the interaction between Group I mGluRs at the single molecule level *in vivo* utilizing our SPReAD technique.

BIOGRAPHICAL SKETCH

Alexander Van Slyke was born in Rochester, New York where his father trained as a medical resident and his mother worked as a special education teacher. Alex graduated from Cooperstown High School. He was awarded the Rensselaer Polytechnic Institute Math and Science Medal upon graduation and proceeded to study Physics for two years at RPI. Seeking a scenic setting and more exposure to the arts, Alex transferred to SUNY Geneseo where he graduated magna cum laude with a Bachelor of Arts in Physics and Mathematics. While there, he was introduced to optical system and circuit design and built a cavity ring-down spectroscopy setup to study the optical scattering properties of atmospheric pollutants. His curiosity and desire to use his recently obtained knowledge of physics and math in a humble effort to bring some good to the world led him to pursue a degree in Biophysics at Cornell University. At Cornell, Alex had the privilege to work under the mentorship of Warren Zipfel. During his tenure there, Alex took advantage of the opportunities available in order to learn a plethora of information about biology, chemistry, physics, microscopy, nanofabrication, and all of the single molecule methods detailed in this dissertation. Upon receiving his Ph.D., Alex will continue on to the University of Pennsylvania to further his quest to serve mankind by applying his knowledge and skills to the field of medical physics.

To my mother

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude and appreciation to my advisor, Professor Warren Zipfel, for all of his guidance and mentorship throughout my graduate education at Cornell University. He introduced me to the fascinating world of single molecule imaging and provided an invaluable guiding hand throughout my graduate studies. He makes himself frequently available to his students, so that brief, unscheduled, insightful meetings were a regular occurrence. This afforded me the freedom to experiment on my own, while having a wise and knowledgeable resource to help decipher results and steer the course of things. I'd like to thank Dr. Rebecca Williams for all of her assistance with confocal imaging techniques and insightful feedback throughout the years. I would like to express my deepest appreciation to my special committee members, Professors Lois Pollack and John Lis, for their scholarly input and helpful guidance on my research projects (and also for putting up with my last-minute notices for meetings and summer talks). I extend a sincere thanks to my undergraduate research advisor, Professor George Marcus, for his role in introducing me to the world of research and sparking my interest in optics.

I'm grateful to Cornell University for providing an environment, rich in opportunities and shared equipment facilities, that broadened my horizons and expertise. Thanks are due to the staff of the Biotechnology Resource Center, especially Carol Bayles and Adam Wojno, for all of the help they provided with FACS. I owe a special thank you to the staff at the Cornell NanoScale Science and Technology Facility for all of their instruction on photolithography techniques and assistance with aspects of fabrication. I had the pleasure of learning the machining trade from Nate Ellis who was both an excellent instructor and a great guy to have a chat with. I'm extremely grateful to the administrative staff of Molecular Biology and Genetics for all of their work keeping the department

running and setting up recruitment and events, as well as some shared laughs. I cannot end this acknowledgement section without expressing my gratitude to the NIH for all of the funding they provided that made my studies possible.

Many thanks go out to the members of the Lis lab for instructing me as I floundered my way from “what’s a gel?” to “sure I can have those plasmids made for you by the end of the week.”

I gratefully acknowledge the assistance from Professor Eric Green and the members of his lab at Columbia University, especially Professor Sy Redding and Dr. Daniel Duzdevich, for their invaluable assistance in getting our “DNA curtains” here at Cornell up and running. Thanks also go out to Dr. Adam Bogdanove and Dr. Fabio Cupri Rinaldi for their assistance with everything TALE related. I also want to acknowledge Professor Paul Kammermeier and Tyler McCullock for their efforts and advice regarding our mGluR project.

I’m extremely grateful to Drs. Avtar Singh and Devin Wakefield for their friendship and all of their assistance and guidance in the ways of graduate school throughout our tenure together. I also had the great pleasure of sharing an office and lab with Juan Wang and Max Kushner who have been helpful in matters related to research, and more importantly, been good friends throughout the years. I want to extend a special thanks to Yi Wen and Marshall Colville who have both been close friends and excellent partners in commiseration throughout this grueling process.

I’d like to acknowledge my dearest friends, Jason Bizzell and Nathan Marucci, for the laughs, the fun, and the camaraderie throughout the years. I’m not sure if I’d have come this far without them.

Last, but certainly not least, I would like to thank my family. I acknowledge my late father for his love, the good times we shared during my childhood, and whose own academic pursuits have inspired mine. I want to express my sincere

gratitude and appreciation to my extended family for their love, emotional support, and gifts. I reserve my penultimate acknowledgement for Asa Bloomingdale, whose love and support has been invaluable, and whose courage has been admirable. Finally, I would like to extend my deepest gratitude to my mother, whose love and devotion has kept me afloat through all the ups and downs in life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PSF	Point Spread Function
NA	Numerical Aperture
AOI	Angle of Incidence
FOV	Field of View
STED	Structured Illumination Microscopy
STORM	Stochastic Optical Reconstruction Microscopy
PALM	Photo-Activation Localization Microscopy
FCS	Fluorescence Correlation Spectroscopy
FCCS	Fluorescence Cross-Correlation Spectroscopy
TIRF	Total Internal Reflection Fluorescence
FRET	Förster Resonance Energy Transfer
FLIM	Fluorescence Lifetime Image Microscopy
FRAP	Fluorescence Recovery After Photobleaching
EMCCD	Electron Multiplying Charge Couple Device
SiM-Pull	Single-Molecule Pull-down
SPReAD	Single Protein Recovery After Dilution
mGluR	Metabotropic Glutamate Receptor
TALE	Transcription Activator-Like Effector
BSA	Bovine Serum Albumin
DNA	Deoxyribonucleic acid
PDMS	Polydimethylsiloxane
PEG	Polyethylene Glycol
CNF	Cornell NanoScale Science & Technology Facility
BRC	Cornell Biotechnology Resource Center
AOTF	Acousto-Optic Tunable Filter
EMSA	Electrophoretic Mobility Shift Assay

CHAPTER 1

SINGLE MOLECULE IMAGING TECHNIQUES: METHODS, PRINCIPLES, AND CONSIDERATIONS

Over the past few decades, single molecule imaging methods have matured from a few groundbreaking techniques to a vast toolkit with applications in nearly every area of biological study. They have provided insights into important individual behaviors, dynamics, and heterogeneities of biomolecules that could not be detected by bulk biochemical or ensemble measurements. The continued development of new techniques and clever combinations of previous techniques have expanded both the scope of their application and the detail that they can provide. Herein, I will review a number of these techniques, the principles that drive them, and considerations that must be taken with experimental design.

INTRODUCTION

Much of our current understanding of biological molecules, structures, interactions, and processes comes from bulk biochemical techniques, ensemble measurements, and genome wide assays [1, 2]. These techniques have given us a detailed understanding of biological function, but there may be transient interactions or critical heterogeneities present that they cannot capture [3].

Fluorescence microscopy and associated fluorescence methods have provided a valuable and reliable toolkit for biophysicists for decades due to their relatively non-invasive nature, capability to visualize processes in real time, and the ability to image multiple different proteins or structures simultaneously by using unique fluorophores on each biomolecule [4].

Single molecule fluorescence imaging has taken these attractive features of fluorescence microscopy and, utilizing the consistent development of new technology and a slew of clever techniques, revolutionized the way we understand biological processes by enabling the investigation of interactions between individual biomolecules.

Over the past few decades, innovation in the field of single molecule studies of biology have produced a wealth of information on the arrangement, dynamics, and forces generated by individual molecules interacting with one another or their substrates. However, the interrogation of biological processes at the single molecule level has inherent issues. These include poor signal-to-noise ratios (SNR) due to the low photon counts produced by single fluorophores, the need to immobilize biomolecules of interest for continued observation of each individual reaction that can create artifact-inducing interactions with the surfaces containing the sample, and the rate of data acquisition [5-8].

A number of different techniques have been developed to study binding kinetics and specificity, dynamics, enzymatic activity, structures, and stoichiometry at the single molecule level, and these will be reviewed in this dissertation, as well as, considerations that must be taken into account when designing and preparing an experiment.

Illumination and Detection

When imaging a single molecule, one can consider it a point source, since the size of a labeled protein, for example, is much smaller than the wavelength of visible light. A single point source will appear as a diffraction pattern on the detector

known as the Airy disk, the properties of which are based on the wavelength of light and the NA of the objective lens [9]. The zeroth-order diffraction spot, called the point spread function (PSF), can be approximated by a Gaussian function with x_o as the location of the peak (the centroid) and *s* as the standard deviation or width of the spot. Abbe's definition of the resolution limit of the microscope is given as *d,* the FWHM of this PSF

$$d = \frac{\lambda}{2NA} \tag{1}$$

with *λ* as the wavelength of light and *NA* as the numerical aperture of the objective lens [10]. However, there is some uncertainty in the location of x_0, due to factors such as the number of photons collected, pixelation as a result of finite detector pixel size, and the signal background. The relationship is described as

$$\langle(\Delta x_0)^2\rangle = \frac{s^2 + a^2/12}{N} + \frac{8\pi s^4 b^2}{a^2 N^2} \tag{2}$$

with Δx_0 as the uncertainty in the centroid location, *s* is the width (standard deviation) of the PSF, *N* is the number of photons collected, *a* is the physical size of the detector pixel, and *b* is the background [8]. For sufficiently large N (~10000), (2) can be estimated as

$$\Delta x_0 = \frac{s}{\sqrt{N}} \tag{3}$$

[11].

Improved detector technologies, such as EMCCD and sCMOS cameras, have enabled the detection of photon counts emitted by a single fluorophore in the millisecond timescale regime and with increasingly small pixel sizes [12]. At

extremely low levels of signal, the background plays a larger role in localization uncertainty, even going so far as to obscure the fluorescence of the desired molecule [6]. An illumination technique that has revolutionized the field of single molecule imaging is total internal fluorescence microscopy (TIRFM) [13]. This utilizes the phenomenon of total internal reflection, where an incident beam of light approaching at, or above, the critical angle reflects off the boundary with a lower index of refraction (i.e. glass into water) with 100% efficiency. However, inside the lower index medium an evanescent wave is generated. An evanescent wave is a surface propagated electromagnetic field whose intensity decays exponentially with distance from the boundary. This decay rate is dependent on the indices of refraction of the two media, and for the purposes of TIRFM, the magnitude of the field intensity is only sufficient to illuminate a 100 nm section above the interface. The original TIRFM setup utilized a prism with an external illumination source to generate the evanescent wave while the emitted fluorescence was collected using an objective lens [13]. A further development that has simplified the implementation of a TIRFM imaging system, called objective TIRF, generates the evanescent field using a laser exiting the objective lens at, or above, the critical angle [14]. The limitations of this technique include the requirement of a high NA oil immersion lens in order to reach such an extreme beam angle and the presence of fringing in the illumination pattern due to constructive and destructive interference occurring from back reflections in the beam path. The latter has been addressed by moving the beam around the perimeter of the back aperture of the objective lens at a higher frequency than the frame rate of the camera resulting in

a uniformly illuminated field due to averaging over all of the diffraction patterns [15].

Surface Passivation and Tethering

The method used to tether molecules to the surface and the prevention of nonspecific surface adsorption are two intertwined concerns. The surface needs to have the ability to secure the molecules of interest in a controlled fashion, yet be sufficiently passivated so that there isn't nonspecific adsorption to the surface [16]. Two commonly used methods of preventing nonspecific surface interactions are the application of a polyethylene glycol (PEG) coating and the use of casein or bovine serum albumin (BSA) to block surfaces that would otherwise adsorb either antibodies or the molecule being studied [17-19]. The drawback is that the molecule being studied is held in close proximity to a surface that may be hydrophobic and/or charged. These conditions may interfere with protein interactions including those with their ligands, substrates, or other proteins [20].

A surface passivation method worth noting, for its unique properties and relevance to my research, is the formation of supported lipid bilayers [21]. These bilayers resemble physiological surfaces and the use of zwitterionic lipids in the bilayer has been shown to prevent the nonspecific adsorption of DNA and proteins [22]. However, while molecules tethered to the bilayer will be confined to the surface, the lipids are also free to diffuse throughout the 2-D bilayer at 1 μm/s allowing molecules tethered to them to move in response to applied force or simple thermal motion [20, 23]. If the desired experiment requires the long-term imaging of a stationary molecule (e.g. stepwise photobleaching), observation of dynamics

of interacting molecules, or stochastic super resolution methods, bilayers may not be the ideal choice of surface passivation or additional measures need to be taken to immobilize the molecule in two dimensions [24] . To this end, continuous bilayers cannot be formed across a discontinuity in the surface allowing lipids to be confined to selected areas [25, 26].

Tethering of molecules is often carried out using either the well-characterized and robust biotin-streptavidin interaction or an antibody-epitope interaction [27-29]. In the specific case of DNA, pH dependent tethering of the DNA molecule itself to a hydrophobic surface can occur [30]. The tethering anchor locations may be randomly distributed by: sparsely distributing streptavidin across the surface prior to passivation, doping the PEGylation mixture with a small concentration of biotin-PEG, or including lipids with a biotinylated head group in a bilayer [20, 31]. The anchors may also be arranged in a predetermined pattern by depositing nanodots or transferring protein onto the surface using a PDMS stamp [32-34].

Data Throughput

The unique ability of single molecule methods to study individual molecules to observe transient behavior or heterogeneities in a population is simultaneously one of its limiting factors [35]. In order to collect data sets large enough to draw statistically sound conclusions, many individual experiments must be observed. Force spectroscopy methods are especially sensitive to the differences in applied force or recorded force across many molecules [36]. High throughput techniques that employ carefully designed external fields and regularly spaced molecules

have been developed to combat this issue [33, 35]. Methods utilizing imaging without concern for the application of precise force lend themselves to parallelization as the extension of molecules is sufficient and uniform illumination across a field of view is more readily accomplished. The issues here lie with orientation and uniformity of alignment and have again been addressed using nanofabricated patterns [34, 37].

SINGLE MOLECULE IMAGING TECHNIQUES

Localization microscopy

Localization microscopy is a technique that enables the location of a fluorophore down to a single nanometer (Figure 1.1a) [7, 8]. The equation and considerations involved in determining the actual location of the fluorescent molecule are given above (Equation 2). This technique can be used for single particle tracking where the location of a particle is recorded in each frame of a movie and, therefore, may be tracked over time. This can also be used for fluorescence colocalization where the frequency of coregistration of fluorophores with different emission spectra is measured and intermolecular interactions are inferred [38, 39]. Furthermore, localization microscopy is the basis of super resolution methods such as Photo-activated Localization Microscopy (PALM) and Stochastic Optical Reconstruction Microscopy (STORM) which enable adjacent molecules to be identified and localized despite having overlapping PSFs using a variety of increasingly clever techniques [40, 41]. These methods are beyond the scope of this document, so I direct you to two detailed reviews on the subject [4, 42].

Single particle tracking can record the dynamics of biomolecules moving through a sample by fitting the PSF of fluorescent puncta in each frame of a movie and tracking the mean squared displacement (MSD) over time (Figure 1.1b) [43, 44]. Insights into the movement of a molecule can help determine whether it is freely diffusing in 1, 2, or 3 dimensions, fixed to a structure or track, or being actively trafficked [39, 45, 46].

Fluorescence colocalization is a technique which compares the location of different structures or molecules in order to infer the presence and frequency of interaction (Figure 1.1c) [47]. The degree of colocalization can be characterized by multiple parameters on either a pixel-by-pixel basis, where signal overlap is measured, or an object basis, where object locations are given by fitted PSF centroids and occupancy within a fixed radius is measured [38, 48, 49]. The experimental setup for fluorescence colocalization experiments is fairly straightforward and requires: two detectors to collect the emission by utilizing different spectral filters with minimal spectral bleed through (ideally none) and molecules of interest labeled with two different colored fluorophores [47]. For widefield imaging, a single camera can be used with an optical splitter that images two different spectral regions to two distinct regions on the camera sensor. The difficulty in colocalization measurements is in the analysis of the data, as a number of factors may incorrectly contribute to the measured degree of colocalization. Numerous analysis methods and software packages are available for the extraction of meaningful conclusions from the data and several methods for dealing with issues such as poor SNR and heterogeneous

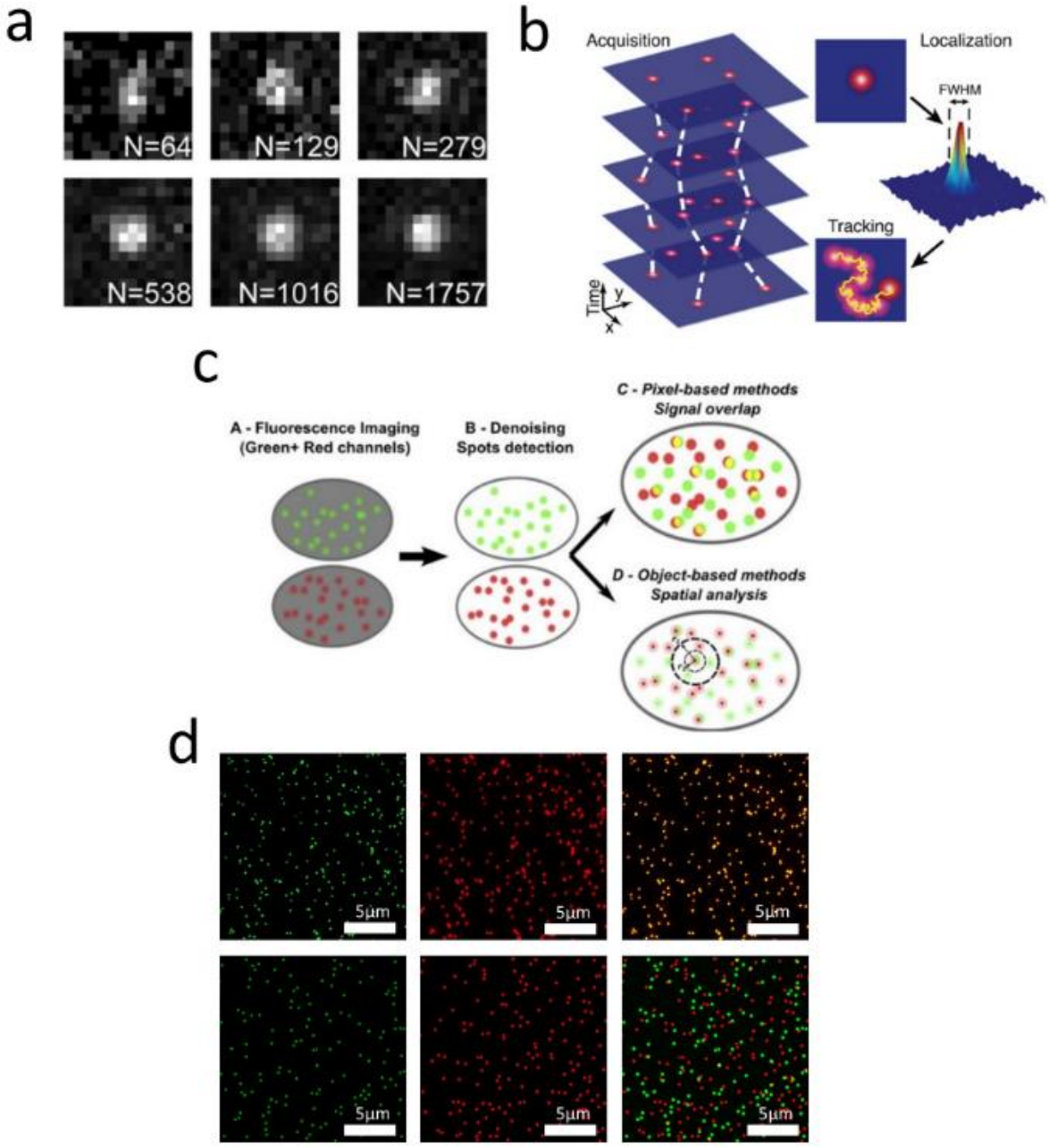


Figure 1.1 Single Molecule Localization Techniques
(a) Images of a fluorescent bead with different photon counts (N) collected. The image quality improves with increasing photons gathered (adapted from Thompson *et al* [8]). (b) Schematic of a particle tracking experiment. The particle is localized in each frame of a movie and the movement is recorded over time (adapted from Manzo *et al* [50]). (c) Workflow of fluorescence colocalization data analysis. Noise is removed prior to either object- or pixel-based colocalization measurements (adapted from *Lagache et al* [49]). (d) Single channel and overlay images of two samples demonstrating 84% (upper panel) and 4% (lower panel) colocalization.

background have been developed over the past few decades [51-53].

Colocalization can be broken down into correlation and co-occurrence, each of which characterize a different aspect. The two primary metrics of colocalization on a pixel basis are Pearson's correlation coefficient (PCC) and Mander's co-occurrence and overlap coefficients (MOC). Calculating both of these values is ideal when attempting to thoroughly characterize interactions. PCC is sensitive to SNR but insensitive to areas of heterogeneous background, while the Mander's method is sensitive to heterogeneous background but insensitive to SNR [51]. Object centroid based colocalization is well suited to handling diffraction limited point sources, such as individual proteins or complexes, even in the presence of noise (Figure 1.1d) [54]. Object methods utilize a marked point process where each spot is considered to be a one dimensional point located at the centroid of each of the fitted fluorescent puncta. Measurements of the mean nearest neighbor distance and the frequency at which the distance to nearest neighbor is less than a user-chosen threshold, usually around 100-200 nm, are used to determine the interaction between the labeled molecules [49, 55].

Stepwise photobleaching

Stepwise photobleaching is a method used to count the number of fluorescently labeled subunits present in a diffraction-limited complex. It is performed by acquiring a movie of a field of sparsely distributed fluorescent molecules at a high laser power, locating the molecules with the same spot-finding algorithms used in the methods discussed above and then counting the number of "steps down" in the fluorescence intensity time traces as the molecule is

photobleached (Figure 1.2a). When the intensity in the N x N sized pixel region reaches the level of the background, the number of steps is inferred to be the number of fluorescent labels present in that spot [53, 54]. Histograms of these step counts taken under different conditions can be assembled to measure changes in protein complex stoichiometry (Figure 1.2b).

Like fluorescence colocalization, the experimental setup for stepwise photobleaching is straightforward, but the analysis presents numerous difficulties [56]. Manually scoring each photobleach trace individually is both time consuming and may differ between individuals because the choice of which traces to score and what constitutes a photobleaching event are not always clear. Issues commonly encountered that complicate the scoring of bleach trace data include: an exponentially decaying background due to gradual bleaching of the environment, varying bleach step height, fluorophore blinking, and the stochastic nature of photobleaching events which can result in bleaching events occurring nearly simultaneously [56]. These attributes of a beaching time trace make developing a reliable method of automated trace scoring particularly difficult. Software packages and algorithms have been developed for this purpose, but dealing with the complexities listed above necessitates that care be taken to avoid skewing the data to fit the hypothesis [57, 58]. Known stoichiometric standards (e.g. a dimer of fluorescent proteins) can be used to test the analysis parameters of your model system and fluorophore [58-60]. Analysis methods will be discussed in Appendix 2 where I compare a published software package to bleach step counting software developed in our laboratory.

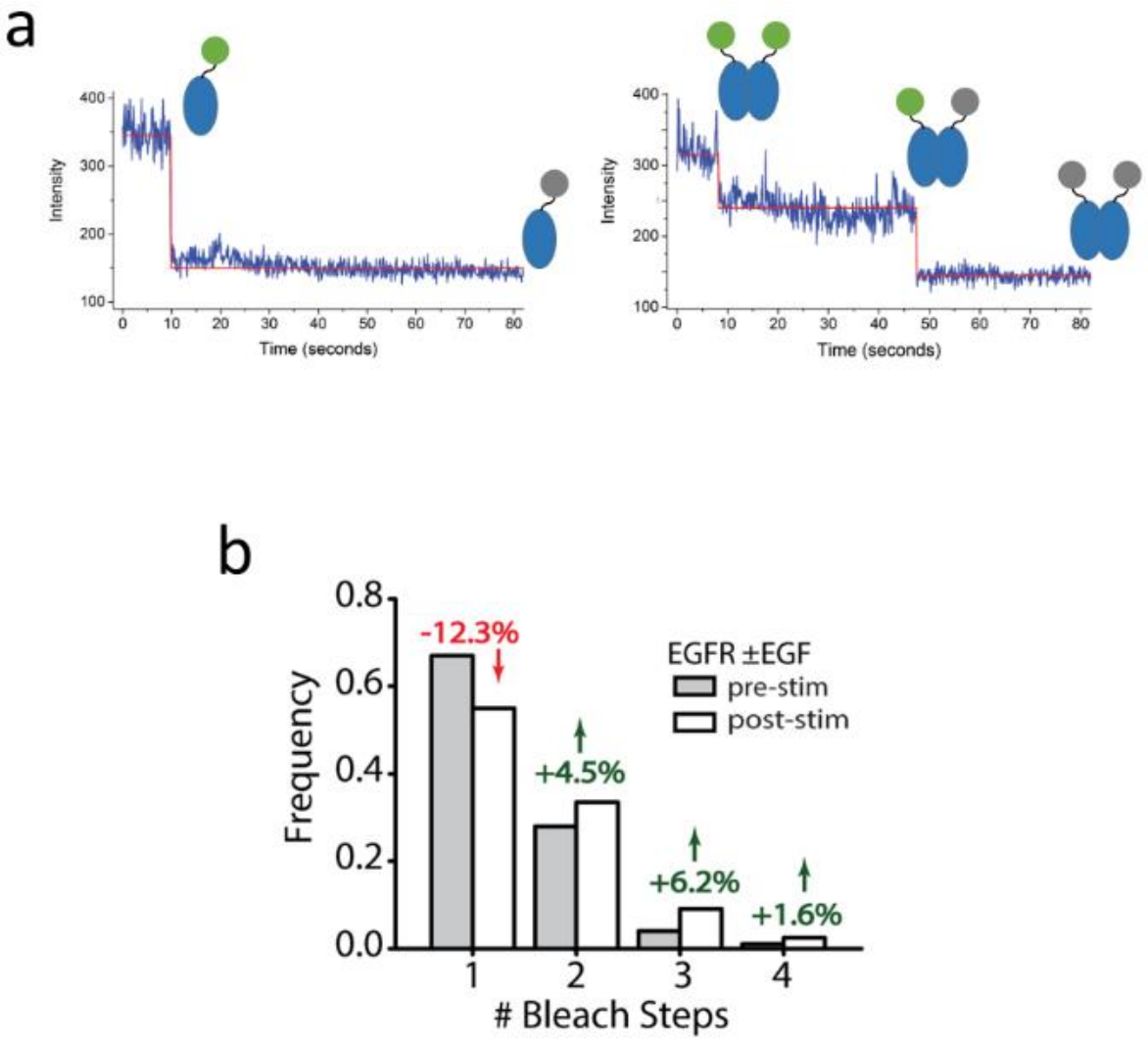


Figure 1.2 Stepwise Photobleaching
(a) Stepwise photobleaching traces with associated diagrams showing one (left) and two (right) bleaching events and their associated protein stoichiometry. Green orbs represent a fluorophore attached to a protein while gray orbs denote that the fluorophore has been photobleached. (b) A histogram of the number of steps per trace for EGFR, showing a shift to larger step counts, and therefore complex stoichiometries, after stimulation with EGF.

Förster resonance energy transfer (FRET)

Förster resonance energy transfer (FRET) is an energy transfer phenomenon which can be used to measure distances down to the nanometer level [61]. FRET occurs when a fluorophore with a shorter emission wavelength (the donor) is excited by an incoming photon, but instead of emitting a fluorescence photon the energy is transferred to a nearby molecule (the acceptor) via dipole-dipole coupling (Figure 1.3a). The coupled molecule then emits a photon corresponding to the acceptor emission spectrum [16, 62]. The efficiency of this energy transfer falls off as a sixth power of the distance between donor and acceptor [62]. As a result, FRET is a sensitive means of measuring distances on the molecular scale

$$R_o = \left[\phi_F \kappa^2 \left(\frac{9000(\ln 10)}{128\pi^5 N_A n^4}\right) J(\lambda)\right]^{\frac{1}{6}} = 0.2108\left(\phi_F \kappa^2 n^{-4} J(\lambda)\right)^{\frac{1}{6}} \quad (4)$$

where R_o is the distance between the donor and acceptor where energy transfer efficiency is 0.5, ϕ_F is the donor fluorescence quantum yield in the absence of the acceptor, κ^2 is the dipole orientation factor, N_A is Avogadro's number (units = number/mole), n is the index of refraction of the medium between the dipoles, and J(λ) is the overlap integral calculated as shown below (units = $M^{-1}\,cm^{-1}\,nm^4$)

$$J(\lambda) = \frac{\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda}{\int_0^\infty F_D(\lambda) d\lambda} \quad (5)$$

FRET occurs at distances up to approximately 10 nm depending on the properties of the donors and acceptors used. One can infer that, if FRET is occurring, then the donor and acceptor molecules are in close proximity and may be interacting in

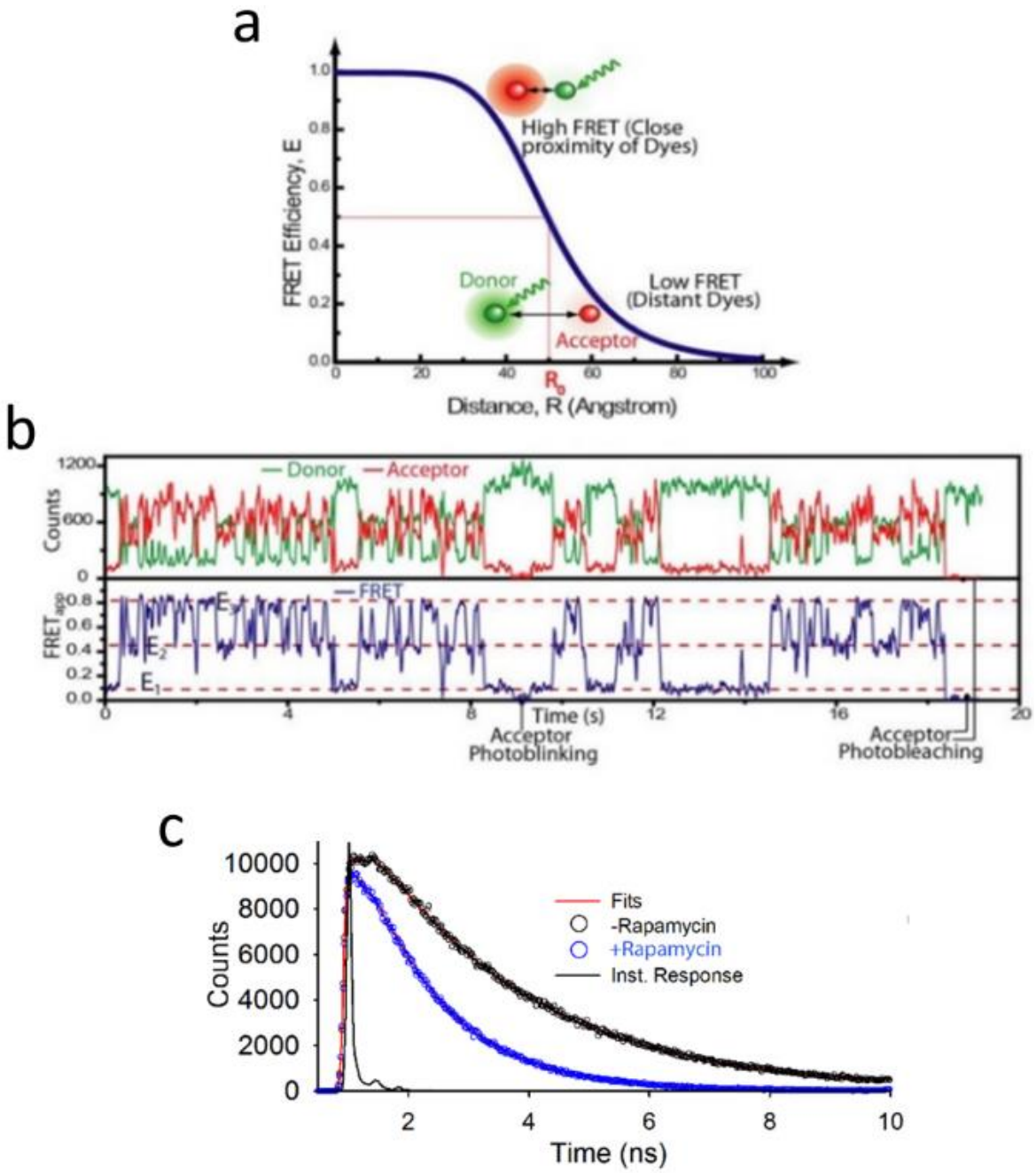


Figure 1.3 Förster Resonance Energy Transfer and Fluorescence Lifetime Imaging (a) Plot of FRET efficiency vs distance between donor and acceptor probes. Efficiency decreases as the probes are moved apart (adapted from *Roy et al* [16]). (b) Example of FRET data between two mobile probes showing the intensity in each color and the FRET efficiency between them (adapted from *Roy et al* [16]). (c) FLIM data of FRB-Venus and FKBP-Cerulean demonstrating the reduction in fluorescence lifetime due to FRET upon the induction of dimerization with rapamycin.

some way [63]. Single particle, or single molecule, FRET (spFRET or smFRET) has been enabled by advancements in detector technology and imaging modalities. This technique measures the energy transfer between a single donor and single acceptor, as opposed to ensemble measurements made using traditional FRET on a sample containing many fluorophores [64]. By extending FRET to the single molecule level, both intramolecular conformational changes and intermolecular interactions, such as ligand binding and complex stoichiometry, can be examined (Figure 1.3b) [63]. Careful placement of the donor and acceptor at specific locations on a protein or protein complex enables conformational changes to be tracked as the distances change. Using FRET, molecular scale conformation changes can be inferred from changes in FRET efficiency. This method has been used to study GTPase induced structural changes in RAF, polymorphism of actin filaments, and helicase activity to name only a few [65, 66]. Intermolecular interactions can also be tracked by labeling either a homogeneous or heterogeneous population of proteins with donor and acceptor chromophores. Typical examples of spFRET are studies of EGFR on the cell membrane or investigations of multiple ligand binding to receptors [67, 68]. spFRET offers the ability to gain insight into the heterogeneity of complexes formed from a population [69]. The relative amounts of donor-acceptor complex, individual donor, and individual acceptor subspecies present in a cell can be ascertained by two color imaging [70]. A more recent application has been the use of piconewton force sensors which contain a donor and acceptor separated by a linker with a known spring constant. Since the efficiency of energy transfer is related to inter-probe

distance, and the spring constant is known, the amount of force applied can be inferred from the transfer efficiency [71, 72].

Fluorescence lifetime imaging microscopy (FLIM)

Fluorescence lifetime imaging microscopy (FLIM) measures changes in the characteristic fluorescence lifetime of a particular fluorophore due to the environment (Figure 1.3c). The fluorescence lifetime is a measure of the time a molecule spends in the excited state after the absorption of a photon and values typically range from 0.2 to 20 ns. Environmental factors such as pH, reactive oxygen species (ROS) induced quenching, or FRET with a nearby acceptor can be quantitatively detected [73-75]. FLIM can be used with FRET to detect energy transfer using only one detector by measuring changes in the lifetime of the donor [76, 77]. When an acceptor molecule is close enough to the donor to have FRET occur, the fluorescence lifetime is decreased because the excited electron(s) is (are) losing their energy to both fluorescence and energy transfer. FLIM has also been used to separate fluorophores with overlapping spectra provided they have unique lifetimes. FLIM can be carried out either in the time domain or frequency domain. Time domain FLIM is performed using a pulsed illumination source in the 10-100 MHz regime so that the fluorescence decay can be gathered between pulses in the absence of illumination [77]. Two-photon microscopy is now typically used with FLIM; both require pulsed illumination sources, and the wavelength of light used for excitation interacts less with objects in the cellular environment [78]. Examples of FLIM applications include the use of lifetime-quantifiable pH sensors, studies of vesicle fusion, measurements of the production of ROS in cancer cells,

and the interaction of flavonols within the nucleus [73, 74, 77, 78].

Fluorescence Recovery After Photobleaching (FRAP)

Fluorescence Recovery After Photobleaching (FRAP) (also known as Fluorescence Photobleaching Recovery (FPR) at Cornell for historical reasons) measures the diffusion rate of fluorescently tagged molecules in a sample by photobleaching a region of the field of view via exposure to a high laser intensity and then recording the time required for unbleached fluorescent molecules to diffuse into the area (Figure 1.4a) [79]. FRAP is of particular usefulness in determining the rate of lateral diffusion of fluorescently labeled biomolecules in, or interacting with, cellular membranes (Figure 1.4b) [80, 81]. FRAP on molecules in solution, or intracellular proteins, can also yield information on binding kinetics and protein motility [82]. These measurements determine whether the motion follows a model of free diffusion, such as a freely mobile protein, or barrier limited diffusion, as is the case with proteins that transiently bind to cellular structures. In addition, in some cases, it can provide reaction kinetics if biochemical reactions must occur before release [83, 84]. Both of these more complex interactions will contribute to the measured recovery time, so models have been generated to fit a number of states that can occur between “reaction dominant” and “pure diffusion” [85]. Multiple data sets in different conditions known to correspond to each of these extremes can be used to generate a more accurate model of both the diffusion rate and binding kinetics and their aggregate effects on the observed recovery [82].

Fluorescence loss in photobleaching (FLIP) is a related technique in which an area is photobleached repeatedly between imaging, unlike FRAP which only bleaches

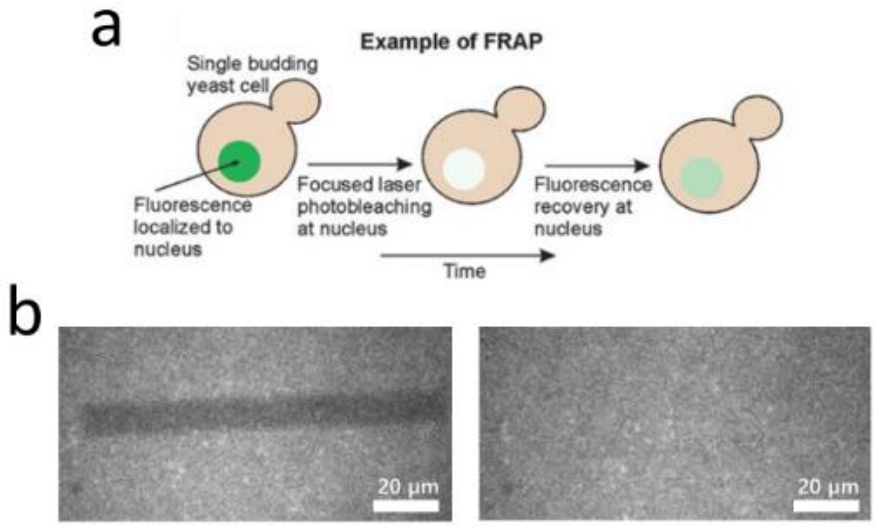


Figure 1.4 Fluorescence Recovery After Photobleaching
(a) Steps in a FRAP experiment of a fluorescently stained nucleus. Molecules are photobleached using high laser intensity. The timescale of fluorescence recovery yields information regarding the mobility of the molecules in the photobleached volume (adapted from *Shashkova et al* [3]). (b) Photobleaching experiment done on a supported lipid bilayer containing 0.1% Lissamine Rhodamine-DOPE. A region showing a photobleached rectangle (left panel) and the same region after fluorescence recovery (right panel).

once before imaging. The measured quantity is the loss of fluorescence in the area surrounding the bleached region. FLIP has applications in studying exchange between compartments and the continuity of organelle membranes [86, 87].

Fluorescence correlation spectroscopy (FCS)

Fluorescence correlation spectroscopy (FCS) and fluorescence cross correlation spectroscopy (FCCS) are techniques that use the temporal correlation of thermodynamic concentration fluctuations of fluorescent molecules in solution to obtain several useful molecular parameters (Figure 1.5a) [88]. By illuminating a small volume on the order of 1 cubic μm (1 fL) and calculating the time autocorrelation function of the intensity fluctuations of the molecules passing through it, concentration, diffusion rate, and reaction kinetics can be calculated (Figure 1.5b) [88, 89]. The number of fluorescently labeled proteins present in a single diffusing complex can be calculated by dividing the total emitted intensity by the number of complexes present in the focal volume at that time (Figure 1.5c). Collection of fluorescence from only the desired excitation volume is critical, so this technique is best suited for confocal or two-photon microscopy [90]. Two color fluorescence cross correlation spectroscopy (FCCS) and FCS-FRET can be used to determine the interaction between two differently labeled molecules by recording the correlation between concentration fluctuations of each color, or the FRET occurring between fluorophores in the sample volume, respectively (Figure 1.5d) [91, 92]. More recently, techniques have been developed to enable FCS to be performed across large fields of view such as scanning FCS (sFCS), TIRF-FCS, and selective plane illumination microscopy (SPIM)-FCS [4].

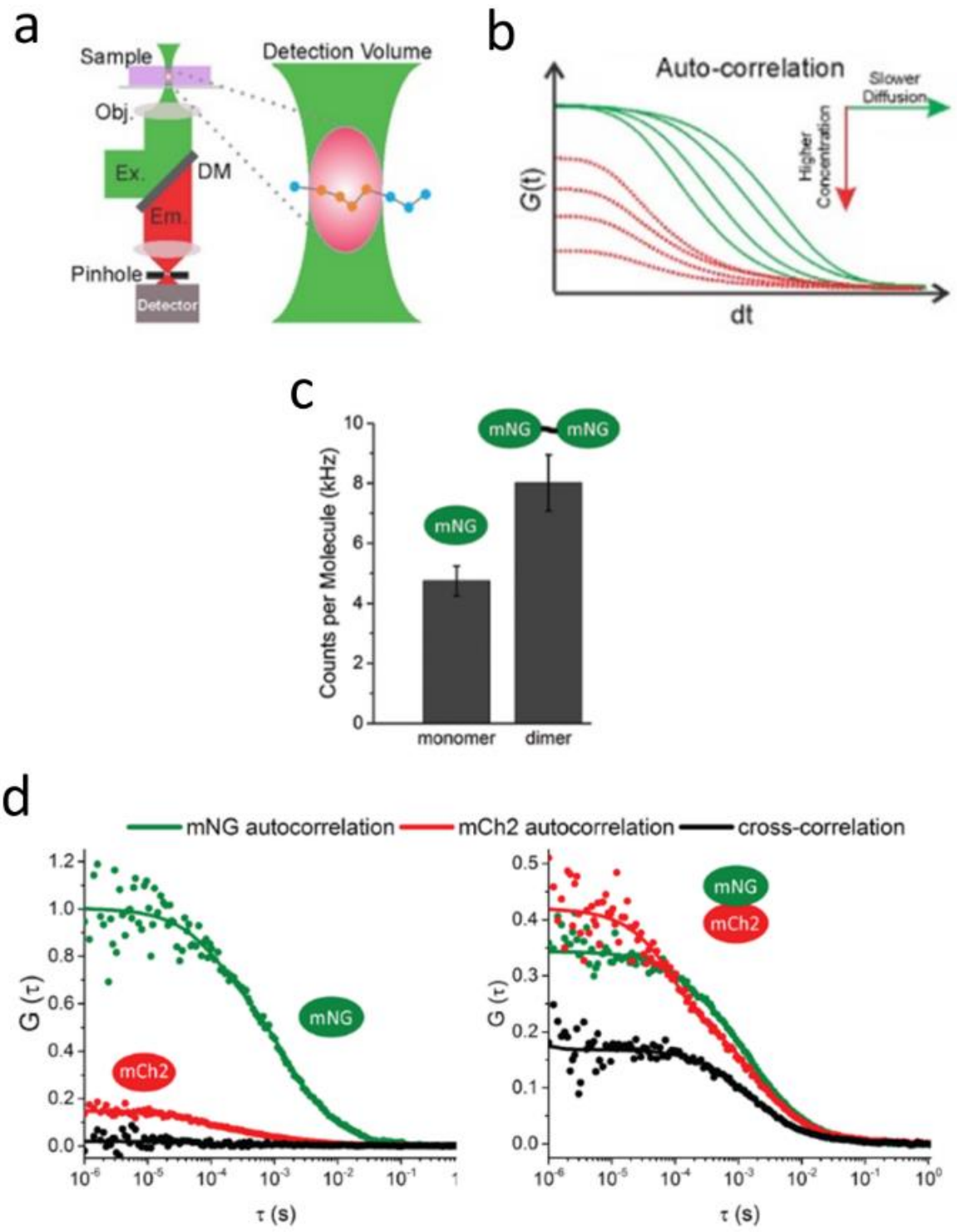


Figure 1.5 Fluorescence Correlation and Cross-Correlation Spectroscopy
(a) Diagram of an FCS experimental setup with a molecule diffusing through the detection volume (adapted from *Liu et al* [4]). (b) Auto-correlation plot that displays information regarding the number of molecules diffusing at different rates (adapted from *Liu et al* [4]). (c) Stoichiometry of diffusing molecules can be calculated from the counts per molecule of a labeled protein. (d) FCCS data showing cross-correlation between a covalent dimer of fluorophores that isn't present when they are not linked.

FCS and FCCS require low concentrations (on the order of tens of nM) of fluorescent molecules so that the illuminated volume contains very few of them at any given time [93]. In this case, the Poissonian concentration fluctuations dominate over photon shot noise and high quality FCS and FCCS data can be obtained. This poses an issue for *in vivo* studies because cells usually express proteins at much higher concentrations, necessitating some form of sample preparation to dilute the native concentration down to FCS-compatible levels. Dilution methods will be discussed throughout this dissertation, as we have developed a novel technique that enables single molecule measurements, including FCS, to be performed *in vivo*.

SINGLE MOLECULE METHODS TO STUDY DNA

There are two primary categories of single molecule techniques used to probe DNA: those that precisely apply and measure forces, such as optical tweezers, magnetic tweezers, and atomic force microscopy (AFM), and those that extend molecules solely for the purpose of imaging. The latter includes methods such as chromatin combing, chromatin elongation by confinement, and microfluidic devices including tethered particle motion and “DNA curtains” [94-97]. The force spectroscopy techniques that I will explore in this chapter are optical and magnetic tweezers because they have led to some important discoveries in the field of biophysics, but I will more broadly detail the imaging methods, as DNA curtains is the technique I chose to utilize in my research.

Force spectroscopy

Optical tweezers refer to a set of techniques which utilize a dielectric material, such

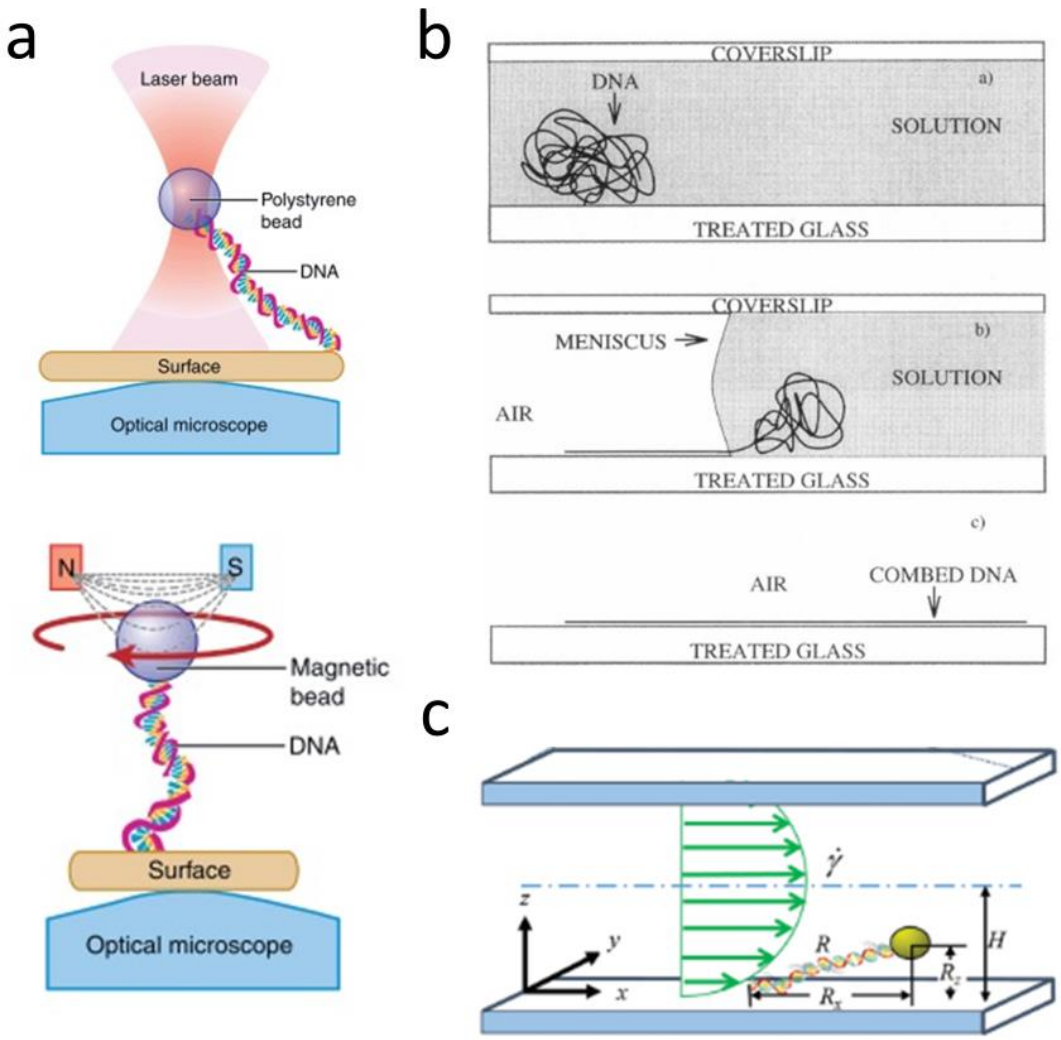


Figure 1.6 Single Molecule DNA Manipulation and Imaging Methods
(a) Diagram of a DNA molecule being extended using optical tweezers (above) and twisted using magnetic tweezers (below) (adapted from *Neuman et al* [98]). (b) DNA being extended on glass using the molecular combing method. The retreating meniscus of DNA solution causes the DNA strands to be outstretched when they adsorb to the surface (adapted from *Allemand et al* [30]) (c) Schematic of the extension of DNA using hydrodynamic force (adapted from *Nir et al* [99]).

as a polystyrene bead, attached to DNA (Figure 1.6a upper panel) [100]. The steep electric field gradient created at the center of a focused laser beam, in combination with the conservation of momentum of photons as they refract in the dielectric material, create an “optical trap” that is used to immobilize a dielectric bead in the center of the beam waist in all 3 dimensions [101]. Force exerted by motor proteins on their substrate can be measured by recording the deflection in the optical trap or by increasing the force applied opposite the motor proteins motion until it stalls [102]. This technique allowed decades-old questions to be answered in a way that was not previously possible. Uncovering the step size and force applied by kinesin and myosin as they move along their substrate, microtubules or actin filaments respectively, are two early examples which demonstrate the power of this technique [102].

In the realm of protein-DNA interaction studies, the force generated by RNA polymerase and the relationship between DNA force and extension [103, 104] were some flagship experiments performed using optical tweezers.

Magnetic tweezers are the second set of single molecule force measurement tools that will be discussed. A magnetic dipole is generated in micron-sized superparamagnetic beads covalently attached to a DNA molecule by the presence of an external magnetic field provided by a rare earth magnet or electromagnet (Figure 1.6a lower panel) [105, 106]. Torsional forces are applied to the bead by manipulating the external magnetic field. These forces can be utilized to supercoil DNA in either handedness, or relax the supercoiling altogether, allowing the DNA configuration itself, or proteins interacting with the DNA, to be

observed. Measurements made using magnetic tweezers, later reinforced by those using optical tweezers, are responsible for confirming that the wormlike chain model, not the freely jointed chain model, accurately models the force-extension relationship for DNA [104, 107, 108]. Some of the protein-DNA interactions examined using magnetic tweezers include the supercoiling preferences and action of topoisomerases [109, 110].

DNA Imaging Methods

Imaging is often combined with the above methods to visualize the behavior of the biomolecules involved while in the presence of external forces [97, 101]. However, some techniques eschew the precision in force application and measurement in order to carry out imaging experiments where force does not need to be quantified [3, 111-113].

Molecular combing involves applying a DNA containing droplet to a substrate, floating an untreated coverslip on top to spread it out, and then evaporating the liquid to apply tension, via the receding air-water interface, leaving the DNA extended in parallel in the imaging plane with the extremities immobilized on the surface (Figure 1.6b) [114]. The force applied by the receding meniscus has been shown to be uniform across the device and independent of the length of the DNA being outstretched [30]. This technique is limited to a narrow pH range depending on the surface involved, preferring pH ~5.5 for hydrophobic surfaces that may alter the behavior of biological molecules [30]. Rehydration of the DNA after combing may result in minor resorption, but does not preclude experiments done in aqueous conditions [30]. Dynamic molecular combing utilizes the combing

of chromatin combined with the use of fluorescence in situ hybridization (FISH) probes to optically map microdeletions and other mutations involved in disease pathology [115]. Recent advancements have extended the range of DNA lengths able to be combed to 10 megabases [116]. These techniques have been used to study replication and transcription, as well as mutations leading to tuberous sclerosis [115, 117-119]. However, molecular combing does not discriminate for the orientation of DNA, resulting in the fragments being aligned either parallel or antiparallel. This is an important consideration when directionality is critical to the phenomenon being studied. Furthermore, the molecules are distributed randomly so particle tracking must be done on an individual basis [94].

Tethered particle motion is a technique which tracks the motion of a bead attached to DNA tethered to the surface of the flow cell in order to measure changes to DNA length in the absence of external force [120]. This allows the DNA to undergo natural conformational changes due to the action of the flow without outside perturbation. This method has been used to gain insights on the protein-DNA interactions of nucleases, recombinases, polymerases, and transcriptional repressors such as LacI [34, 121-125]. A multiplexed variation was developed using a nanofabricated master to cast PDMS stamps, which were then coated with a neutravidin solution and stamped onto epoxydized glass. These neutravidin patches would tether biotinylated DNA to the surface in a regularly spaced grid pattern allowing numerous particle tracking experiments to be performed in a single field of view [34].

The use of hydrodynamic force to extend molecules has been utilized in a

number of different methods. In general, DNA is tethered to the surface by one end and continuous buffer flow applies a hydrodynamic force to the DNA fragment and serves to stretch it in the direction of flow (Figure 1.6c). This technique has been used to investigate the dynamics of replication and replisome formation, to elucidate how the DNA repair protein RecA locates its target, and to test models of the diffusion constant for proteins sliding along the helical backbone of DNA [126-130]. The effects of this hydrodynamic force on the protein-DNA interaction may create a bias for the direction of motion or artificially inflate the k_D by encouraging the factors to dissociate from the DNA [131]. In order to observe outstretched DNA molecules in the absence of force, some implementations of this technique use DNA tethered at both ends, either with the same functional group, which loses directionality, or with a second moiety [132-135].

DNA Curtains

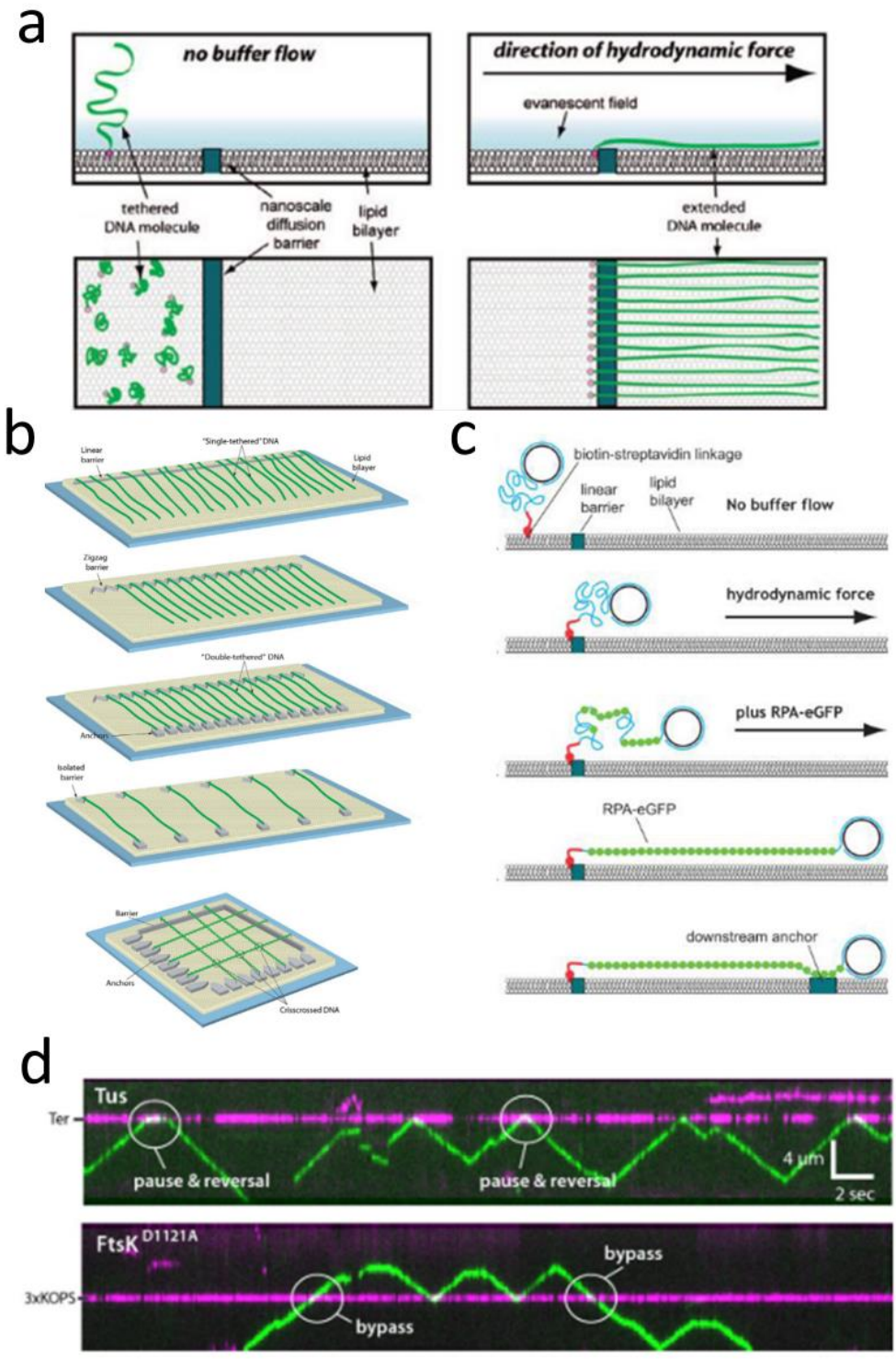
One hydrodynamic extension technique that stands out is the “DNA curtains” platform [37]. In brief, this technique involves a nanopatterned flow cell coated in a lipid bilayer and the use of buffer flow to exert a hydrodynamic force on tethered DNA molecules in the flow cell. This force both extends DNA molecules in the imaging plane and allows reaction conditions to be changed rapidly and without interrupting imaging acquisition (Figure 1.7a). Dozens or hundreds of DNA molecules are extended in parallel, aligned at their ends, and can be observed in a single field of view. This enables the gathering of a statistically relevant number of replicate experiments in a short period of time, overcoming the low throughput barrier common to single molecule methods.

The use of supported lipid bilayers is made possible by the “curtain rod” diffusion barriers, thus the high throughput nature of flow cell based single molecule imaging experiments can still be harnessed while preserving the advantage of using bilayers to passivate the surface. There have been numerous improvements or alterations made to the geometry of the flow cell, materials used, method of fabrication, and source of the DNA used to form the curtains.

DNA Curtain Geometry

The diffusion barriers in the initial DNA curtains flow cells were simply scratches etched in the glass by hand using a diamond scribe. These formed irregular patterns that varied in width from 1-5um [20]. In order to have a narrow, uniform “curtain rod” to hang the DNA curtains from, the next development replaced the irregular glass scratch barriers with nanopatterned chrome rods deposited on the surface (Figure 1.7b top image) [136]. These curtain rods were 100 nm wide and 30 nm tall to minimize interactions between biomolecules and the chromium and keep the DNA extended in the TIRF illumination field. A variation on this straight curtain rod pattern is a sawtooth shape which at low concentrations of DNA results in parallel DNA molecules spaced a fixed distance apart from one another (Figure 1.7b second image) [137]. This removes any effect that a crowded environment may have on the interaction being observed. The next major innovation was the development of “double tethered” DNA curtains [134]. These utilize a pentagon shaped chrome pedestal spaced 12 microns away from the curtain rod to anchor the non-biotinylated end of the DNA (Figure 1.7b third image). The second tethering reaction is mediated by antibodies deposited on

Figure 1.7 Schematic of Several DNA Curtains Implementations
(a) Side and top schematic of the DNA curtains flow cell. DNA freely diffuses through the bilayer in the absence of force (left panels) but is moved to the nanobarriers and outstretched in the presence of hydrodynamic force (right panels) (adapted from *Fazio et al* [136]). (b) Five types of curtain nanopatterns used for different types of protein-DNA experiments (adapted from *Collins et al* [24]) (c) Formation of single stranded DNA curtains using rolling circle amplification and replication protein A to linearize the DNA (adapted from *Gibb et al* [138]). (d) Interaction of a DNA translocase, FtsK, with other proteins bound to DNA. Details of dynamics such as passing and pausing events can be seen using DNA curtains (adapted from *Lee et al* [139])


a
no buffer flow
direction of hydrodynamic force
evanescent field
tethered DNA molecule
nanoscale diffusion barrier
lipid bilayer
extended DNA molecule
b
"Single-tethered" DNA
Linear barrier
Lipid bilayer
Zigzag barrier
"Double-tethered" DNA
Anchors
Isolated barrier
Barrier
Anchors
Crisscrossed DNA
c
biotin-streptavidin linkage
linear barrier
lipid bilayer
No buffer flow
hydrodynamic force
plus RPA-eGFP
RPA-eGFP
downstream anchor
d
Tus
Ter
pause & reversal
pause & reversal
4 µm
2 sec
FtsK D1121A
3xKOPS
bypass
bypass

the pedestal and a digoxigenin attached to the end of the DNA opposite the biotinylation site. Further developments in the geometry of the flow cell include crisscrossed DNA and isolated DNA tethers (Figure 1.7b fourth and bottom images) [140, 141].

Materials and Methods for Fabrication

Nanopatterned diffusion barriers provide excellent control over the shape and dimensions of the curtain rod, and subsequently, on the DNA curtain itself. Chromium has been the most common material used to form the diffusion barriers, but hydrogen silsesquioxane has also been shown to be a valid alternative [136, 142]. The first patterned flow cells were produced using electron-beam lithography. The extraordinary resolution and precision afforded by this technique enabled chromium curtain rod designs as narrow as 100 nm to be deposited successfully [136]. However, the tools required to perform this type of lithography are not widely available to researchers and the cost in both fabrication time and materials involved to produce a single device is significant. In an effort to reduce the amount of time required to produce a single device, and to minimize the amount of electron-beam access required, a method for producing the devices using nanoimprint lithography was developed [143]. Photolithography provides a cheaper, faster, and more widely available method of fabrication than the electron-beam alternative, with a loss in resolution that, while important in the sub 10nm feature sized field of semiconductors, is minor on the 100 nm scale of the curtain rods. A method using photolithography to produce a SU-8 master for casting of PDMS devices in bulk has greatly streamlined the production process, providing that such a device is

compatible with the materials involved in a given experiment [144].

DNA Substrates

Canonical DNA curtains utilize lambda phage DNA as it provides a substrate that is well-studied, large enough to be able to visualize motion along the DNA strand, and when linearized, is flanked with 12 bp overhangs which enable ligation of different functional groups to either end of the DNA molecule [145]. Furthermore, cloning target sites into the lambda phage is a well-developed technique, providing flexibility in the types of DNA-protein interactions that can be examined [146].

However, many important and interesting biological processes such as replication initiation and recombination involve single stranded DNA (ssDNA). To study these, a method of forming ssDNA curtains was developed using rolling circle amplification (RCA) (Figure 1.7c) [138, 147, 148].

Uses of DNA Curtains

The DNA curtains platform provides a powerful tool in the single molecule protein-DNA imaging toolkit. By providing hundreds of identical experiments in a single field of view, generating statistically relevant conclusions, once the bottleneck of single molecule throughput, can be drawn in a fraction of the time. Using DNA curtains, it was discovered that the presence of a protospacer adjacent motif (PAM) was important at all stages of CRISPR-Cas9 function [149]. Studies of the related Cascade system in *E. coli* demonstrated that Cascade could bind to target sequences without the presence of a PAM, but recruitment and function of the Cas3 nuclease did not occur on sites lacking a PAM [150]. The ability to rapidly

change the buffer composition while continuously imaging has provided a means to thoroughly study the interaction of ATPases with their DNA substrate. The Msh2-Msh6 complex was discovered to search for mismatches via 1D diffusion, but the dissociation of the complex from DNA after repair was mediated by ADP/ATP exchange [132]. The heterotrimeric translocase RecBCD was shown to proceed with enough force to push or eject many DNA binding proteins including LacI, dEcoRI, stalled elongation complex, and RNAP [151, 152]. Further studies with another DNA translocase, FtsK, demonstrated that it can also push or evict DNA- bound proteins, but RecBCD was found to exert enough force to overwhelm FtsK (Figure 1.7d) [139]. The ability to work with single stranded DNA has elucidated the role of replication protein A in preventing spontaneous annealing of dsDNA breaks [153]. Presynaptic complex assembly, homologous recombination, and meiosis specific recombination were also examined using ssDNA curtains [154-158]. These are only some examples of how the various geometries and techniques of the DNA curtains platform have been used to further our understanding of protein-DNA interaction at the single molecule level.

SINGLE MOLECULE STOICHIOMETRY MEASUREMENTS

Studying the binding or other enzymatic actions of purified proteins *in vitro* can provide valuable information about their function and kinetics, but the gold standard for functional biological studies on the cellular level is to capture this behavior as it occurs *in vivo*. However, constitutively expressed proteins in cells are present in concentrations much higher than the upper limit for single molecule measurements, so some steps must be taken to achieve a sparsely populated field

of view [159].

Single-molecule Pull-down

One particularly effective method of achieving this sparse fluorophore density is Single-Molecule Pull-down (SiM-Pull) which utilizes a microfluidic device coated with a biotin doped layer of PEG [31]. Biontinylated antibodies are attached to the surface via the biotin-streptavidin interaction. Cell lysate is introduced to the flow cell and the antibodies selectively “pull down” their epitope, often a FLAG tag or other commonly used moiety, immobilizing the protein of interest and any factors with which it was interacting (Figure 1.8a) [160]. Fluorescent labels permit the investigation of protein interaction or complex stoichiometry using stepwise photobleaching, FRET, or fluorescence colocalization [31, 161]. This technique has been used in studies of receptor complex formation and behavior such as the dependence of cancer on growth-factor receptor stoichiometry and phosphorylation of G-protein coupled receptors (GPCRs) [161-163]. The concentration of antibody injected, fraction of biotin in the PEGylation mixture, or amount of cell lysate added can be adjusted to achieve a field of view with the desired density of labeled protein [160]. One method of note that has built upon this technique is the use of shallow flow chambers and on-chip lysis to capture the approximate spatial distribution of proteins in the cell prior to lysis [164].

While these methods permit proteins to be endogenously expressed at physiological concentrations prior to imaging, thereby promoting interactions formed by second-order binding processes that may otherwise be hampered by underexpression, the use of detergents and cell lysis may disrupt the integrity of

some protein complexes, leading to false negatives or altered stoichiometry (evidence shown in Chapter 3) [167].

Inducible Promoters

Another method for achieving the requisite sparse density of fluorophores is the use of an inducible promoter, such as the Tet-On/Tet-Off system, to manipulate the level of protein expression [168]. Some other examples of inducible promoters, which I will not address in this dissertation, utilize RNA microinjections, heavy metal ions, heat shock, or hormones to induce gene expression [168-170]. The tetracycline inducible system can be designed to activate transcription (Tet-On) or to repress transcription (Tet-Off) in the presence of doxycycline (Figure 1.8b) [171]. Both systems use a *trans*-activator (tTA or rtTA) consisting of a fusion between the Tet repressor of the Tn*10* Tc resistance operon from *E. coli* and a portion of VP16, a strong transactivator from the herpes simplex virus. These transactivators bind to copies of the *tet* operon (*tetO)* upstream of a minimal promoter such as the cytomegalovirus (CMV) promoter to induce transcription [171]. The difference between the two is the absence of 4 amino acids in the Tet-On reverse *trans*-activator (rtTA) which causes the protein to bind the operon in the presence of doxycycline (Figure 1.8c,d).

These systems are advantageous because they allow transcription of a particular gene to occur only when specified by the experimentalist, especially the Tet systems as they have been shown to have a very low level of leaky expression (Figure 1.4c,d) [168]. This control over expression can be used to activate a process prior to, or during an experiment, to acutely examine the effect of the

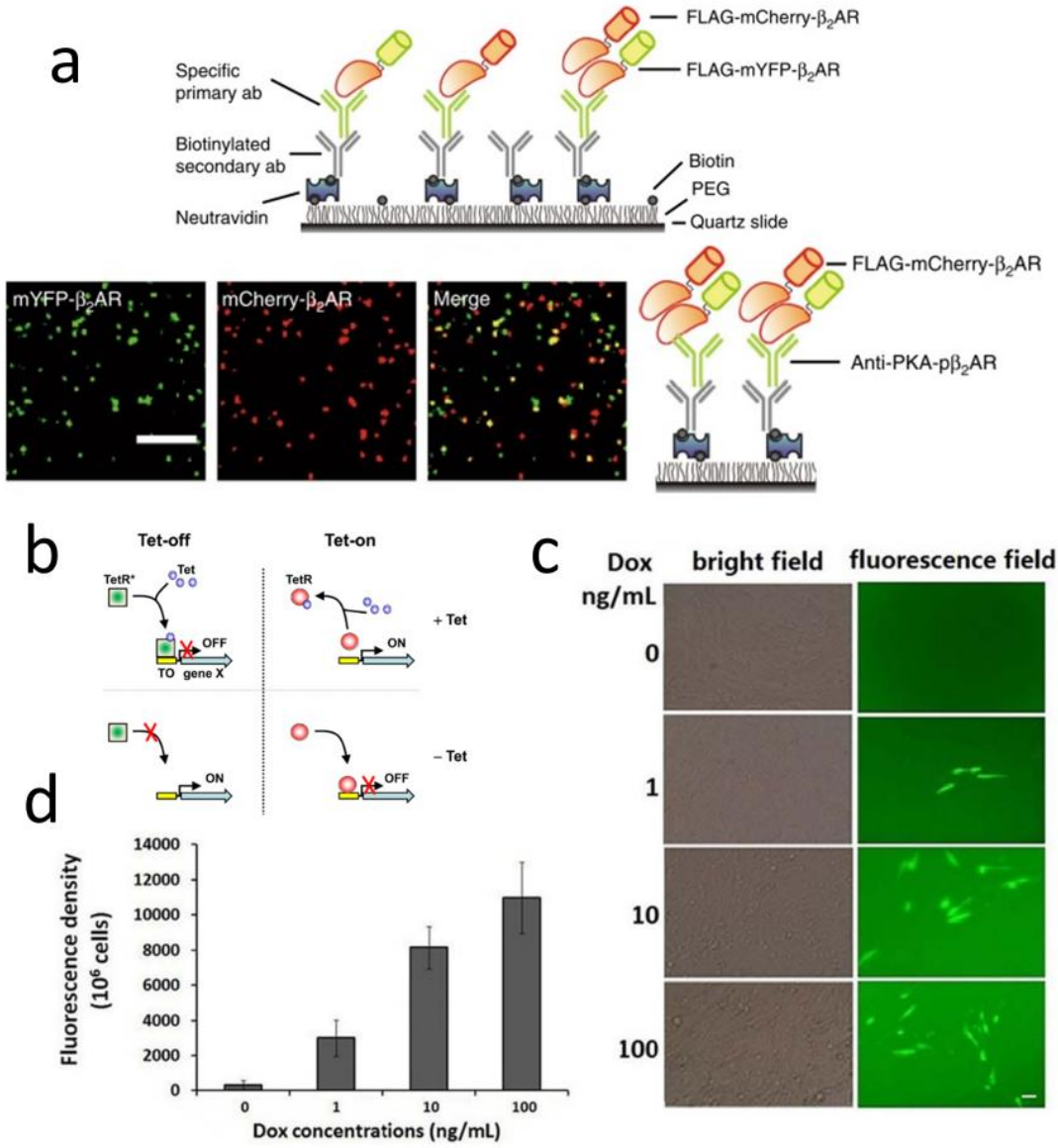


Figure 1.8 Sample Preparation Methods for Imaging Proteins Expressed *in vivo*
(a) Layout of a Single Molecule Pull-down flow cell with PEGylated surfaces and antibodies attached to the surface. Colocalization between the immunoprecipitated proteins can be seen as overlapping colors in the fluorescence images (adapted from *Shen et al* [161]). (b) Schematic of Tet-on and Tet-off systems showing the binding of TetR or TetR* in the presence or absence of tetracycline (adapted from *Gomez-Martinez et al* [165]). (c) Images of cells expressing PAC1-YFP under a Tet inducible promoter at varying levels doxycycline. (d) Quantitative analysis of the images in panel c showing an increase in fluorescence with increasing doxycycline concentration (panels c and d adapted from *Rongjie et al* [166]).

transcribed gene or to minimize the deleterious effects of a protein that is toxic to the cell. Another advantage of the Tet-On system is the ability to induce low levels of expression by adjusting the concentration of doxycycline or the timing of induction relative to imaging [166].

Inducible promoters provide a dilute concentration of proteins to be imaged *in vivo* and thus permit imaging to be done in the absence of detergents or other harsh denaturing conditions used during lysis [172]. However, the low level of expression reduces the frequency of intermolecular collisions, so it may preclude or reduce formation of complexes formed by interactions after reaching the membrane.

REFERENCES

1. Gilmour, D.S. and J.T. Lis, *In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster.* Mol Cell Biol, 1985. **5**(8): p. 2009-18.
2. Gingras, A.C., et al., *Analysis of protein complexes using mass spectrometry.* Nat Rev Mol Cell Biol, 2007. **8**(8): p. 645-54.
3. Shashkova, S. and M.C. Leake, *Single-molecule fluorescence microscopy review: shedding new light on old problems.* Biosci Rep, 2017. **37**(4).
4. Liu, Z., L.D. Lavis, and E. Betzig, *Imaging live-cell dynamics and structure at the single-molecule level.* Mol Cell, 2015. **58**(4): p. 644-59.
5. Michalet, X., et al., *Detectors for single-molecule fluorescence imaging and spectroscopy.* J Mod Opt, 2007. **54**(2-3): p. 239.
6. Michalet, X., et al., *Development of new photon-counting detectors for single-molecule fluorescence microscopy.* Philos Trans R Soc Lond B Biol Sci, 2013. **368**(1611): p. 20120035.
7. Bobroff, N., *Position Measurement With a Resolution and Noise-Limited Instrument.* Rev Sci Instrum, 1986. **57**(6): p. 1152-1157.
8. Thompson, R.E., D.R. Larson, and W.W. Webb, *Precise nanometer localization analysis for individual fluorescent probes.* Biophys J, 2002. **82**(5): p. 2775-83.
9. Born, M. and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. 7 ed. 1999, Cambridge: Cambridge University Press.

10. Mortensen, K.I., et al., *Optimized localization analysis for single-molecule tracking and super-resolution microscopy.* Nat Methods, 2010. **7**(5): p. 377-81.

11. Sauer, M., *Localization microscopy coming of age: from concepts to biological impact.* J Cell Sci, 2013. **126**(Pt 16): p. 3505-13.

12. Michalet, X., et al., *New photon-counting detectors for single-molecule fluorescence spectroscopy and imaging.* Proc SPIE Int Soc Opt Eng, 2011. **8033**: p. 803316.

13. Axelrod, D., *Cell-substrate contacts illuminated by total internal reflection fluorescence.* J Cell Biol, 1981. **89**(1): p. 141-5.

14. Stout, A.L. and D. Axelrod, *Evanescent field excitation of fluorescence by epi-illumination microscopy.* Appl Opt, 1989. **28**(24): p. 5237-42.

15. Colville, M.J., et al., *High-speed device synchronization in optical microscopy with an open-source hardware control platform.* bioRxiv, 2019: p. 533349.

16. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nat Methods, 2008. **5**(6): p. 507-16.

17. Sofia, S.J., V.V. Premnath, and E.W. Merrill, *Poly(ethylene oxide) Grafted to Silicon Surfaces: Grafting Density and Protein Adsorption.* Macromolecules, 1998. **31**(15): p. 5059-70.

18. Heyes, C.D., et al., *Synthesis, patterning and applications of star-shaped poly(ethylene glycol) biofunctionalized surfaces.* Mol Biosyst, 2007. **3**(6): p. 419-30.

19. Greene, E.C. and K. Mizuuchi, *Direct observation of single MuB polymers: evidence for a DNA-dependent conformational change for generating an active target complex.* Mol Cell, 2002. **9**(5): p. 1079-89.

20. Graneli, A., et al., *Organized arrays of individual DNA molecules tethered to supported lipid bilayers.* Langmuir, 2006. **22**(1): p. 292-9.

21. Sackmann, E., *Supported membranes: scientific and practical applications.* Science, 1996. **271**(5245): p. 43-8.

22. Glasmastar, K., et al., *Protein adsorption on supported phospholipid bilayers.* J Colloid Interface Sci, 2002. **246**(1): p. 40-7.

23. Yoshina-Ishii, C. and S.G. Boxer, *Arrays of mobile tethered vesicles on supported lipid bilayers.* J Am Chem Soc, 2003. **125**(13): p. 3696-7.

24. Collins, B.E., et al., *DNA curtains: novel tools for imaging protein-nucleic acid interactions at the single-molecule level.* Methods Cell Biol, 2014. **123**: p. 217-34.

25. Groves, J.T. and S.G. Boxer, *Micropattern formation in supported lipid membranes.* Acc Chem Res, 2002. **35**(3): p. 149-57.

26. Cremer, P. and S.G. Boxer, *Formation and Spreading of Lipid Bilayers on Planar Glass Supports.* J. Phys. Chem. B., 1999. **103**(13): p. 2554-2559.

27. Green, N.M., *Avidin. 1. The Use of (14-C)Biotin for Kinetic Studies and for Assay.* Biochem J, 1963. **89**: p. 585-91.

28. Hart, S.M. and C. Basu, *Optimization of a digoxigenin-based immunoassay system for gene detection in Arabidopsis thaliana.* J Biomol Tech, 2009. **20**(2): p. 96-100.

29. Hofmann, K. and Y. Kiso, *An approach to the targeted attachment of peptides and proteins to solid supports.* Proc Natl Acad Sci U S A, 1976. **73**(10): p. 3516-8.

30. Allemand, J.F., et al., *pH-dependent specific binding and combing of DNA.* Biophys J, 1997. **73**(4): p. 2064-70.

31. Jain, A., et al., *Probing cellular protein complexes using single-molecule pull-down.* Nature, 2011. **473**(7348): p. 484-8.

32. Palma, M., et al., *Selective biomolecular nanoarrays for parallel single-molecule investigations.* J Am Chem Soc, 2011. **133**(20): p. 7656-9.

33. De Vlaminck, I., et al., *Highly parallel magnetic tweezers by targeted DNA tethering.* Nano Lett, 2011. **11**(12): p. 5489-93.

34. Plenat, T., et al., *High-throughput single-molecule analysis of DNA-protein interactions by tethered particle motion.* Nucleic Acids Res, 2012. **40**(12): p. e89.

35. Soltani, M., et al., *Nanophotonic trapping for precise manipulation of biomolecular arrays.* Nat Nanotechnol, 2014. **9**(6): p. 448-52.

36. Ribeck, N. and O.A. Saleh, *Multiplexed single-molecule measurements with magnetic tweezers.* Rev Sci Instrum, 2008. **79**(9): p. 094301.

37. Greene, E.C., et al., *DNA curtains for high-throughput single-molecule optical imaging.* Methods Enzymol, 2010. **472**: p. 293-315.

38. Manders, E.M.M., F.J. Verbeek, and J.A. Aten, *Measurement of co-localization of objects in dual-colour confocal images.* Journal of Microscopy, 1993. **169**: p. 375-382.

39. Yildiz, A. and P.R. Selvin, *Fluorescence imaging with one nanometer accuracy: application to molecular motors.* Acc Chem Res, 2005. **38**(7): p. 574-82.

40. Betzig, E., et al., *Imaging intracellular fluorescent proteins at nanometer resolution.* Science, 2006. **313**(5793): p. 1642-5.

41. Rust, M.J., M. Bates, and X. Zhuang, *Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).* Nat Methods, 2006. **3**(10): p. 793-5.

42. Thorley, J.A., J. Pike, and J.Z. Rappoport, *Chapter 14 - Super-resolution Microscopy: A Comparison of Commercially Available Options*, in *Fluorescence Microscopy*, A. Cornea and P.M. Conn, Editors. 2014, Academic Press: Boston. p. 199-212.

43. Barak, L.S. and W.W. Webb, *Fluorescent low density lipoprotein for observation of dynamics of individual receptor complexes on cultured human fibroblasts.* J Cell Biol, 1981. **90**(3): p. 595-604.

44. Levi, V. and E. Gratton, *Exploring dynamics in living cells by tracking single particles.* Cell Biochem Biophys, 2007. **48**(1): p. 1-15.

45. Lakadamyali, M., et al., *Visualizing infection of individual influenza viruses.* Proc Natl Acad Sci U S A, 2003. **100**(16): p. 9280-5.

46. Sokoll, S., et al., *Fast Three-Dimensional Single-Particle Tracking in Natural Brain Tissue.* Biophys J, 2015. **109**(7): p. 1463-71.

47. Smallcombe, A., *Multicolor imaging: the important question of co-localization.* Biotechniques, 2001. **30**(6): p. 1240-2, 1244-6.

48. Pearson, K., *Mathematical contributions to the theory of evolution III. Regression, heredity and panmixia.* Philos Trans R Soc Lond B Biol Sci, 1896. **187**: p. 253-318.

49. Lagache, T., et al., *Statistical analysis of molecule colocalization in bioimaging.* Cytometry A, 2015. **87**(6): p. 568-79.

50. Manzo, C. and M.F. Garcia-Parajo, *A review of progress in single particle tracking: from methods to biophysical insights.* Rep Prog Phys, 2015. **78**(12): p. 124601.

51. Aaron, J.S., A.B. Taylor, and T.L. Chew, *Image co-localization - co-occurrence versus correlation.* J Cell Sci, 2018. **131**(3).

52. Dunn, K.W., M.M. Kamocka, and J.H. McDonald, *A practical guide to evaluating colocalization in biological microscopy.* Am J Physiol Cell Physiol, 2011. **300**(4): p. C723-42.

53. Zinchuk, V. and O. Grossenbacher-Zinchuk, *Quantitative colocalization analysis of fluorescence microscopy images.* Curr Protoc Cell Biol, 2014. **62**: p. Unit 4 19 1-14.

54. Smal, I., et al., *Quantitative comparison of spot detection methods in fluorescence microscopy.* IEEE Trans Med Imaging, 2010. **29**(2): p. 282-301.

55. Clark, P.J. and F.C. Evans, *Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations.* Ecology, 1954. **35**(4): p. 445-453.

56. McGuire, H., et al., *Automating single subunit counting of membrane proteins in mammalian cells.* J Biol Chem, 2012. **287**(43): p. 35912-21.

57. Das, S.K., et al., *Membrane protein stoichiometry determined from the step-wise photobleaching of dye-labelled subunits.* Chembiochem, 2007. **8**(9): p. 994-9.

58. Ulbrich, M.H. and E.Y. Isacoff, *Subunit counting in membrane-bound proteins.* Nat Methods, 2007. **4**(4): p. 319-21.

59. Blunck, R., et al., *Fluorescence detection of the movement of single KcsA subunits reveals cooperativity.* Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20263-8.

60. Reiner, A., R.J. Arant, and E.Y. Isacoff, *Assembly stoichiometry of the GluK2/GluK5 kainate receptor complex.* Cell Rep, 2012. **1**(3): p. 234-40.

61. Stryer, L., *Fluorescence spectroscopy of proteins.* Science, 1968. **162**(3853): p. 526-33.

62. Forster, T., *Experimental and Theoretical Investigation of the Intermolecular Transfer of Electronic Excitation Energy.* Zeitschrift Naturforsch A, 1949. **4**: p. 321-327.

63. Ha, T., *Single-molecule fluorescence methods for the study of nucleic acids.* Curr Opin Struct Biol, 2001. **11**(3): p. 287-92.

64. Ha, T., et al., *Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor.* Proc Natl Acad Sci U S A, 1996. **93**(13): p. 6264-8.

65. Hibino, K., et al., *A RasGTP-induced conformational change in C-RAF is essential for accurate molecular recognition.* Biophys J, 2009. **97**(5): p. 1277-87.

66. Kozuka, J., et al., *Dynamic polymorphism of single actin molecules in the actin filament.* Nat Chem Biol, 2006. **2**(2): p. 83-6.

67. Sako, Y., S. Minoghchi, and T. Yanagida, *Single-molecule imaging of EGFR signalling on the surface of living cells.* Nat Cell Biol, 2000. **2**(3): p. 168-72.

68. Schutz, G.J., W. Trabesinger, and T. Schmidt, *Direct observation of ligand colocalization on individual receptor molecules.* Biophys J, 1998. **74**(5): p. 2223-6.

69. Ha, T., et al., *Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism.* Proc Natl Acad Sci U S A, 1999. **96**(3): p. 893-8.

70. Nguyen, V.T., Y. Kamio, and H. Higuchi, *Single-molecule imaging of cooperative assembly of gamma-hemolysin on erythrocyte membranes.* EMBO J, 2003. **22**(19): p. 4968-79.

71. Freikamp, A., et al., *Investigating piconewton forces in cells by FRET-based molecular force microscopy.* J Struct Biol, 2017. **197**(1): p. 37-42.

72. Ringer, P., et al., *Multiplexing molecular tension sensors reveals piconewton force gradient across talin-1.* Nat Methods, 2017. **14**(11): p. 1090-1096.

73. Sanders, R., et al., *Quantitative pH imaging in cells using confocal fluorescence lifetime imaging microscopy.* Anal Biochem, 1995. **227**(2): p. 302-8.

74. Sud, D., et al., *Time-resolved optical imaging provides a molecular snapshot of altered metabolic function in living human cancer cell models.* Opt Express, 2006. **14**(10): p. 4412-26.

75. Becker, W., et al., *Fluorescence lifetime images and correlation spectra obtained by multidimensional TCSPC*. Vol. 5700. 2005, Proc of SPIE. 144-151.

76. Chang, C.W., D. Sud, and M.A. Mycek, *Fluorescence Lifetime Imaging Microscopy*, in *Methods in Cell Biology*. 2007, Academic Press. p. 495-524.

77. Oida, T., Y. Sako, and A. Kusumi, *Fluorescence lifetime imaging microscopy (flimscopy). Methodology development and application to studies of endosome fusion in single cells.* Biophys J, 1993. **64**(3): p. 676-85.

78. Mueller-Harvey, I., et al., *Two-photon excitation with pico-second fluorescence lifetime imaging to detect nuclear association of flavanols.* Anal Chim Acta, 2012. **719**: p. 68-75.

79. Axelrod, D., et al., *Mobility measurement by analysis of fluorescence photobleaching recovery kinetics.* Biophys J, 1976. **16**(9): p. 1055-69.

80. Goehring, N.W., et al., *FRAP analysis of membrane-associated proteins: lateral diffusion and membrane-cytoplasmic exchange.* Biophys J, 2010. **99**(8): p. 2443-52.

81. Poo, M. and R.A. Cone, *Lateral diffusion of rhodopsin in the photoreceptor membrane.* Nature, 1974. **247**(5441): p. 438-41.

82. Sadegh Zadeh, K., H.J. Montas, and A. Shirmohammadi, *Identification of biomolecule mass transport and binding rate parameters in living cells by inverse modeling.* Theor Biol Med Model, 2006. **3**: p. 36.

83. Bulinski, J.C., et al., *Rapid dynamics of the microtubule binding of ensconsin in vivo.* J Cell Sci, 2001. **114**(Pt 21): p. 3885-97.

84. Qin, K., et al., *Inactive-state preassembly of G(q)-coupled receptors and G(q) heterotrimers.* Nat Chem Biol, 2011. **7**(10): p. 740-7.

85. Sprague, B.L., et al., *Analysis of binding reactions by fluorescence recovery after photobleaching.* Biophys J, 2004. **86**(6): p. 3473-95.

86. Terasaki, M., *Dynamics of the endoplasmic reticulum and golgi apparatus during early sea urchin development.* Mol Biol Cell, 2000. **11**(3): p. 897-914.

87. Wustner, D., et al., *Quantitative fluorescence loss in photobleaching for analysis of protein transport and aggregation.* BMC Bioinformatics, 2012. **13**: p. 296.

88. Magde, D., E.L. Elson, and W.W. Webb, *Fluorescence correlation spectroscopy. II. An experimental realization.* Biopolymers, 1974. **13**(1): p. 29-61.

89. Magde, D., E. Elson, and W.W. Webb, *Thermodynamic Fluctuations in a Reacting System---Measurement by Fluorescence Correlation Spectroscopy.* Physical Review Letters, 1972. **29**(11): p. 705-708.

90. Berland, K.M., P.T. So, and E. Gratton, *Two-photon fluorescence correlation spectroscopy: method and application to the intracellular environment.* Biophys J, 1995. **68**(2): p. 694-701.

91. Bacia, K. and P. Schwille, *Practical guidelines for dual-color fluorescence cross-correlation spectroscopy.* Nat Protoc, 2007. **2**(11): p. 2842-56.

92. Torres, T. and M. Levitus, *Measuring conformational dynamics: a new FCS-FRET approach.* J Phys Chem B, 2007. **111**(25): p. 7392-400.

93. Hedde, P.N., et al., *Stimulated emission depletion-based raster image correlation spectroscopy reveals biomolecular dynamics in live cells.* Nat Commun, 2013. **4**: p. 2093.

94. Robison, A.D. and I.J. Finkelstein, *High-throughput single-molecule studies of protein-DNA interactions.* FEBS Lett, 2014. **588**(19): p. 3539-46.

95. Bustamante, C., Z. Bryant, and S.B. Smith, *Ten years of tension: single-molecule DNA mechanics.* Nature, 2003. **421**(6921): p. 423-7.

96. Neuman, K.C. and A. Nagy, *Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy.* Nat Methods, 2008. **5**(6): p. 491-505.

97. Iino, R., et al., *Single-molecule imaging and manipulation of biomolecular machines and systems.* Biochim Biophys Acta Gen Subj, 2018. **1862**(2): p. 241-252.

98. Neuman, K.C., T. Lionnet, and J.F. Allemand, *Single-Molecule Micromanipulation Techniques.* Annual Review of Materials Research, 2007. **37**(1): p. 33-67.

99. Nir, G., et al., *The role of near-wall drag effects in the dynamics of tethered DNA under shear flow.* Soft Matter, 2018. **14**(12): p. 2219-2226.

100. Ashkin, A., et al., *Observation of a single-beam gradient force optical trap for dielectric particles.* Opt Lett, 1986. **11**(5): p. 288.

101. Leake, M.C., *The physics of life: one molecule at a time.* Philos Trans R Soc Lond B Biol Sci, 2013. **368**(1611): p. 20120248.

102. Simmons, R., *Molecular motors: single-molecule mechanics.* Curr Biol, 1996. **6**(4): p. 392-4.

103. Yin, H., et al., *Transcription against an applied force.* Science, 1995. **270**(5242): p. 1653-7.

104. Wang, M.D., et al., *Stretching DNA with optical tweezers.* Biophys J, 1997. **72**(3): p. 1335-46.

105. F.H.C., C. and A.F.W. Hughes, *The physical properties of cytoplasm: A study by means of the magnetic particle method Part I. Experimental.* Exp Cell Res, 1950. **1**(1): p. 37-80.

106. De Vlaminck, I. and C. Dekker, *Recent advances in magnetic tweezers.* Annu Rev Biophys, 2012. **41**: p. 453-72.

107. Bustamante, C., et al., *Entropic elasticity of lambda-phage DNA.* Science, 1994. **265**(5178): p. 1599-600.

108. Smith, S.B., L. Finzi, and C. Bustamante, *Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads.* Science, 1992. **258**(5085): p. 1122-6.

109. Crisona, N.J., et al., *Preferential relaxation of positively supercoiled DNA by E. coli topoisomerase IV in single-molecule and ensemble measurements.* Genes Dev, 2000. **14**(22): p. 2881-92.

110. Strick, T.R., V. Croquette, and D. Bensimon, *Single-molecule analysis of DNA uncoiling by a type II topoisomerase.* Nature, 2000. **404**(6780): p. 901-4.

111. Monachino, E., L.M. Spenkelink, and A.M. van Oijen, *Watching cellular machinery in action, one molecule at a time.* J Cell Biol, 2017. **216**(1): p. 41-51.

112. Okamoto, K., M. Hiroshima, and Y. Sako, *Single-molecule fluorescence-based analysis of protein conformation, interaction, and oligomerization in cellular systems.* Biophys Rev, 2018. **10**(2): p. 317-326.

113. Kaur, G., J.S. Lewis, and A.M. van Oijen, *Shining a Spotlight on DNA: Single-Molecule Methods to Visualise DNA.* Molecules, 2019. **24**(3).

114. Bensimon, A., et al., *Alignment and sensitive detection of DNA by a moving interface.* Science, 1994. **265**(5181): p. 2096-8.

115. Michalet, X., et al., *Dynamic molecular combing: stretching the whole human genome for high-resolution studies.* Science, 1997. **277**(5331): p. 1518-23.

116. Kaykov, A., et al., *Molecular Combing of Single DNA Molecules on the 10 Megabase Scale.* Sci Rep, 2016. **6**: p. 19636.

117. Marheineke, K., et al., *Use of DNA combing to study DNA replication in Xenopus and human cell-free systems.* Methods Mol Biol, 2009. **521**: p. 575-603.

118. Herrick, J. and A. Bensimon, *Single molecule analysis of DNA replication.* Biochimie, 1999. **81**(8-9): p. 859-71.

119. Gueroui, Z., et al., *Observation by fluorescence microscopy of transcription on single combed DNA.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 6005-10.

120. Lindner, M., et al., *Force-free measurements of the conformations of DNA molecules tethered to a wall.* Phys Rev E Stat Nonlin Soft Matter Phys, 2011. **83**(1 Pt 1): p. 011916.

121. Dennis, C., et al., *RuvAB-directed branch migration of individual Holliday junctions is impeded by sequence heterology.* EMBO J, 2004. **23**(12): p. 2413-22.

122. Rutkauskas, D., et al., *Tetramer opening in LacI-mediated DNA looping.* Proc Natl Acad Sci U S A, 2009. **106**(39): p. 16627-32.

123. Schafer, D.A., et al., *Transcription by single molecules of RNA polymerase observed by light microscopy.* Nature, 1991. **352**(6334): p. 444-8.

124. Yin, H., R. Landick, and J. Gelles, *Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule.* Biophys J, 1994. **67**(6): p. 2468-78.

125. Zaremba, M., et al., *DNA synapsis through transient tetramerization triggers cleavage by EcI18kI restriction enzyme.* Nucleic Acids Res, 2010. **38**(20): p. 7142-54.

126. Yardimci, H., et al., *Bypass of a protein barrier by a replicative DNA helicase.* Nature, 2012. **492**(7428): p. 205-9.

127. Yao, N.Y., et al., *Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression.* Proc Natl Acad Sci U S A, 2009. **106**(32): p. 13236-41.

128. Bagchi, B., P.C. Blainey, and X.S. Xie, *Diffusion constant of a nonspecifically bound protein undergoing curvilinear motion along DNA.* J Phys Chem B, 2008. **112**(19): p. 6282-4.

129. Loparo, J.J., et al., *Simultaneous single-molecule measurements of phage T7 replisome composition and function reveal the mechanism of polymerase exchange.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3584-9.

130. Ragunathan, K., C. Liu, and T. Ha, *RecA filament sliding on DNA facilitates homology search.* Elife, 2012. **1**: p. e00067.

131. Graneli, A., et al., *Long-distance lateral diffusion of human Rad51 on double-stranded DNA.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1221-6.

132. Gorman, J., et al., *Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6.* Mol Cell, 2007. **28**(3): p. 359-70.

133. Cuculis, L., et al., *Direct observation of TALE protein dynamics reveals a two-state search mechanism.* Nat Commun, 2015. **6**: p. 7277.

134. Gorman, J., et al., *Nanofabricated racks of aligned and anchored DNA substrates for single-molecule imaging.* Langmuir, 2010. **26**(2): p. 1372-9.

135. Hamdan, S.M., et al., *Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis.* Nature, 2009. **457**(7227): p. 336-9.

136. Fazio, T., et al., *DNA curtains and nanoscale curtain rods: high-throughput tools for single molecule imaging.* Langmuir, 2008. **24**(18): p. 10524-31.

137. Visnapuu, M.L., et al., *Parallel arrays of geometric nanowells for assembling curtains of DNA with controlled lateral dispersion.* Langmuir, 2008. **24**(19): p. 11293-9.

138. Gibb, B., et al., *Single-stranded DNA curtains for real-time single-molecule visualization of protein-nucleic acid interactions.* Anal Chem, 2012. **84**(18): p. 7607-12.

139. Lee, J.Y., et al., *Single-molecule imaging of FtsK translocation reveals mechanistic features of protein-protein collisions on DNA.* Mol Cell, 2014. **54**(5): p. 832-43.

140. Gorman, J., et al., *Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair.* Proc Natl Acad Sci U S A, 2012. **109**(45): p. E3074-83.

141. Wang, F., et al., *The promoter-search mechanism of Escherichia coli RNA polymerase is dominated by three-dimensional diffusion.* Nat Struct Mol Biol, 2013. **20**(2): p. 174-81.

142. Fazio, T.A., et al., *Assembly of DNA curtains using hydrogen silsesquioxane as a barrier to lipid diffusion.* Anal Chem, 2012. **84**(18): p. 7613-7.

143. Fazio, T.A., et al., *Fabrication of Nanoscale "Curtain Rods" for DNA Curtains Using Nanoimprint Lithography.* J Vac Sci Technol A, 2009. **27**(6): p. 3095-3098.

144. Robison, A.D. and I.J. Finkelstein, *Rapid prototyping of multichannel microfluidic devices for single-molecule DNA curtain imaging.* Anal Chem, 2014. **86**(9): p. 4157-63.

145. Sanger, F., et al., *Nucleotide sequence of bacteriophage λ DNA.* Journal of Molecular Biology, 1982. **162**(4): p. 729-773.

146. Chauthaiwale, V.M., A. Therwath, and V.V. Deshpande, *Bacteriophage lambda as a cloning vector.* Microbiol Rev, 1992. **56**(4): p. 577-91.

147. Poffenberger, K.L. and B. Roizman, *A noninverting genome of a viable herpes simplex virus 1: presence of head-to-tail linkages in packaged genomes and requirements for circularization after infection.* J Virol, 1985. **53**(2): p. 587-95.

148. Ali, M.M., et al., *Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine.* Chem Soc Rev, 2014. **43**(10): p. 3324-41.

149. Sternberg, S.H., et al., *DNA interrogation by the CRISPR RNA-guided endonuclease Cas9.* Nature, 2014. **507**(7490): p. 62-7.

150. Redding, S., et al., *Surveillance and Processing of Foreign DNA by the Escherichia coli CRISPR-Cas System.* Cell, 2015. **163**(4): p. 854-65.

151. Finkelstein, I.J., M.L. Visnapuu, and E.C. Greene, *Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase.* Nature, 2010. **468**(7326): p. 983-7.

152. Terakawa, T., et al., *Sequential eviction of crowded nucleoprotein complexes by the exonuclease RecBCD molecular motor.* Proc Natl Acad Sci U S A, 2017. **114**(31): p. E6322-E6331.

153. Deng, S.K., et al., *RPA antagonizes microhomology-mediated repair of DNA double-strand breaks.* Nat Struct Mol Biol, 2014. **21**(4): p. 405-12.

154. Gibb, B., et al., *Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging.* PLoS One, 2014. **9**(2): p. e87922.

155. Crickard, J.B., et al., *Regulation of Hed1 and Rad54 binding during maturation of the meiosis-specific presynaptic complex.* EMBO J, 2018. **37**(7).

156. Crickard, J.B., et al., *Spontaneous self-segregation of Rad51 and Dmc1 DNA recombinases within mixed recombinase filaments.* J Biol Chem, 2018. **293**(11): p. 4191-4200.

157. Crickard, J.B., et al., *Dynamic interactions of the homologous pairing 2 (Hop2)-meiotic nuclear divisions 1 (Mnd1) protein complex with meiotic presynaptic filaments in budding yeast.* J Biol Chem, 2019. **294**(2): p. 490-501.

158. Crickard, J.B., et al., *Meiosis-specific recombinase Dmc1 is a potent inhibitor of the Srs2 antirecombinase.* Proc Natl Acad Sci U S A, 2018. **115**(43): p. E10041-E10048.

159. Lewis, J.S., S. Jergic, and N.E. Dixon, *The E. coli DNA Replication Fork.* Enzymes, 2016. **39**: p. 31-88.

160. Jain, A., et al., *Single-molecule pull-down for studying protein interactions.* Nat Protoc, 2012. **7**(3): p. 445-52.

161. Shen, A., et al., *Functionally distinct and selectively phosphorylated GPCR subpopulations co-exist in a single cell.* Nat Commun, 2018. **9**(1): p. 1050.

162. Litwin, D.B., et al., *The structural arrangement at intersubunit interfaces in homomeric kainate receptors.* Sci Rep, 2019. **9**(1): p. 6969.

163. Lee, H.W., et al., *Profiling of protein-protein interactions via single-molecule techniques predicts the dependence of cancers on growth-factor receptors.* Nat Biomed Eng, 2018. **2**(4): p. 239-253.

164. Wang, X., et al., *Toward Single-Cell Single-Molecule Pull-Down.* Biophys J, 2018. **115**(2): p. 283-288.

165. Gomez-Martinez, M., D. Schmitz, and A. Hergovich, *Generation of stable human cell lines with Tetracycline-inducible (Tet-on) shRNA or cDNA expression.* J Vis Exp, 2013(73): p. e50171.

166. Yu, R., et al., *Dimer-Dependent Intrinsic/Basal Activity of the Class B G Protein-Coupled Receptor PAC1 Promotes Cellular Anti-Apoptotic Activity through Wnt/β-Catenin Pathways that Are Associated with Dimer Endocytosis*. Vol. 9. 2014. e113913.

167. Lee, Y.-C., et al., *Impact of Detergents on Membrane Protein Complex Isolation.* Journal of Proteome Research, 2018. **17**(1): p. 348-358.

168. Gossen, M. and H. Bujard, *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.* Proc Natl Acad Sci U S A, 1992. **89**(12): p. 5547-51.

169. Lee, F., et al., *Glucocorticoids regulate expression of dihydrofolate reductase cDNA in mouse mammary tumour virus chimaeric plasmids.* Nature, 1981. **294**(5838): p. 228-32.

170. Mayo, K.E., R. Warren, and R.D. Palmiter, *The mouse metallothionein-I gene is transcriptionally regulated by cadmium following transfection into human or mouse cells.* Cell, 1982. **29**(1): p. 99-108.

171. Baron, U. and H. Bujard, *Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances.* Methods Enzymol, 2000. **327**: p. 401-21.

172. Brown, R.B. and J. Audet, *Current techniques for single-cell lysis.* J R Soc Interface, 2008. **5 Suppl 2**: p. S131-8.

CHAPTER 2

SIMPLIFYING THE FABRICATION OF A HIGH THROUGHPUT MICROFLUIDIC DEVICE TO IMAGE PROTEIN-DNA INTERACTIONS AT THE SINGLE MOLECULE LEVEL

DNA-protein interactions are crucial to the proper function and survival of cells. Proteins with a modular DNA binding domain, such as Transcription Activator-Like Effectors (TALEs) and the CRISPR-Cas9 system, have been utilized over the past decade for gene silencing, insertion, deletion, and visualization. Single molecule imaging methods provide insights into the dynamics and the target search mechanisms of these proteins that would not be possible using ensemble techniques. The "DNA curtains" technique is a microfluidic based single molecule imaging method well-suited to investigate these interactions in a highly multiplexed fashion. We demonstrate that photolithography and subsequent chromium deposition can be used to create the characteristic nanobarriers used in DNA curtains flow cells. As a result, the DNA curtains devices can be produced for a fraction of the cost utilizing equipment that is more accessible to the general scientific community. We demonstrated successful protein-DNA binding within the flow cell using two model proteins, LacI and EcoRI^{E111Q}. However, we determined that Transcription Activator-Like Effectors (TALEs), which are interesting due to their ability to target any DNA sequence, are not compatible with the DNA curtains technique due to their high affinity for the chromium patterns used to immobilize the DNA.

INTRODUCTION

Single molecule studies provide information on the behavior of individual molecules as they interact with their DNA substrate, elucidating dynamics and other transient binding states that would otherwise not have been revealed [3, 4]. However, single molecule imaging methods require the biomolecules to function *in vitro* where the conditions are very different from their native environment [7]. Furthermore, traditional single molecule methods are limited by their low throughput, although multiplexed methods have been developed to address this issue [7-12]. The nanopatterned microfluidic device that forms the basis of the DNA curtains platform enables high throughput observation of proteins interacting with their DNA substrate on the single molecule level (Figure 2.1a) [5]. The use of a supported lipid bilayer with zwitterionic lipid headgroups to passivate the surface provides an environment that more closely resembles the native environment of biomolecules [7]. The current methods of patterning and fabricating DNA curtains flow cells rely on electron-beam lithography to produce ~200 nm features on either the device itself, or a master used for nanoimprint lithography or PDMS casting (Figure 2.1b) [15-17]. “Double tether curtains” expand on the DNA curtains technology by adding 500 x 370 nm^2 chromium nanopillars spaced 13 μm away from the “curtain rod” designed to catch the free end of the DNA molecule using a digoxigenin-label and antibodies deposited on the nanopillars (Figure 2.1c) [18]. By anchoring the DNA at both ends, the molecules remain outstretched in the imaging plane in the absence of flow (Figure 2.1d). We demonstrate that photolithography provides a more cost effective, expedient, and accessible

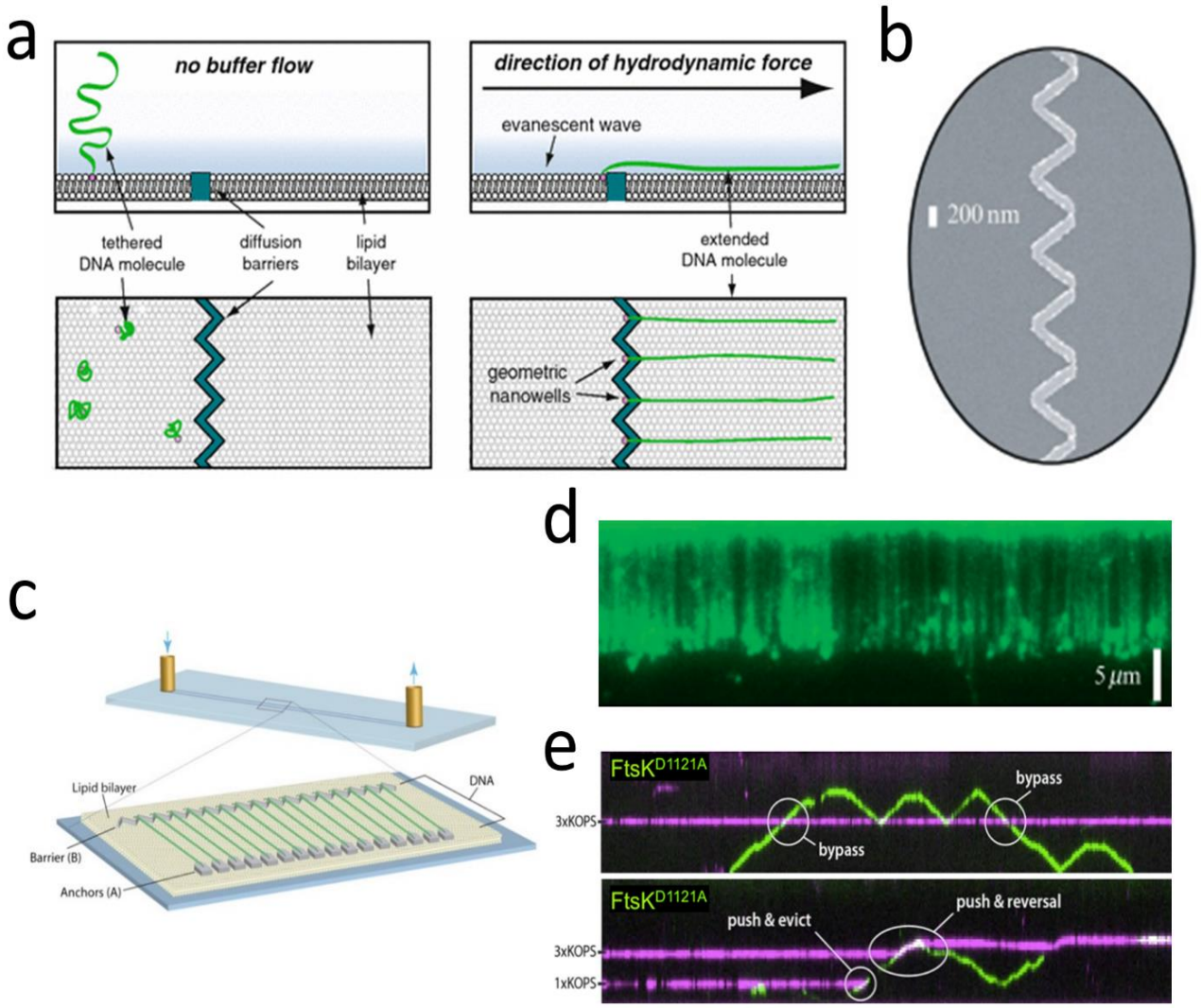


Figure 2.1 Overview of the DNA Curtains Flow Cell
(a) Side and top view of DNA curtains flow cell. DNA attached to the bilayer can freely diffuse in the absence of external forces (left panel) but is pushed up against, and then extended over chromium nanobarriers by a hydrodynamic force (right panel). Adapted from Visnapuu, M *et al.* [5] (b) SEM image of DNA curtains showing the "sawtooth" pattern which encourages spacing between extended DNA molecules. Adapted from Green, E *et al.*[13] (c) The layout of a "double tether" setup, containing a second set of chromium pillars 500 x 370 nm^2 (in the direction of flow and perpendicular to the direction of flow respectively) in order to keep DNA extended in the absence of force. (d) Double tethered YOYO-1 labeled DNA remaining extended in the absence of force (e) FtsK, a translocase moving along extended DNA molecules and colliding with an immobile mutant, $FtsK^{D1121A}$, to provide an example of the dynamics that can be observed using DNA curtains. (Panels d and e adapted from Lee *et al.*[14]).

method of fabricating DNA curtains devices with feature sizes on the order of 500 nm. We further demonstrate the formation of a bilayer in these flow cells and confirm the expected binding behavior of two well-characterized proteins, LacI and EcoRI^{E111Q} [19, 20].

DNA curtains provide an excellent platform to observe the kinetics and dynamics of proteins interacting with their DNA substrates because they provide the ability to easily change the buffer composition, introduce protein while DNA is already under observation, and do so in a high throughput fashion so that statistically sound can be drawn. As one example of this, FtsK, a DNA translocase was observed moving along a DNA molecule and interacting with an immobile FtsK^{D1121A} mutant (Figure 2.1e) [14].

We utilized DNA curtains devices to investigate the search mechanism and dynamics of transcription activator-like effectors (TALEs) due to their unique modular DNA binding properties [21]. Transcription activator-like effectors are a type of protein from *Xanthomonas* which bind to a specific target sequence of DNA in order to induce transcription of specific genes related to bacterial virulence in the host plant cell (Figure 2.2a) [22]. TALEs contain repeats of a 34 amino acid domain which vary from one another in the 12th and 13th residues labeled the repeat variable di-residues (RVDs) (Figure 2.2b) [21]. Their identity determines the affinity of that domain for a specific nucleotide [23]. Designer Transcription Activator-Like Effector (dTALE) can be generated to target any arbitrary DNA sequence by changing the number of domains and the identity of the RVDs [24].

dTALEs enable targeted gene deletion when paired with a nuclease, control

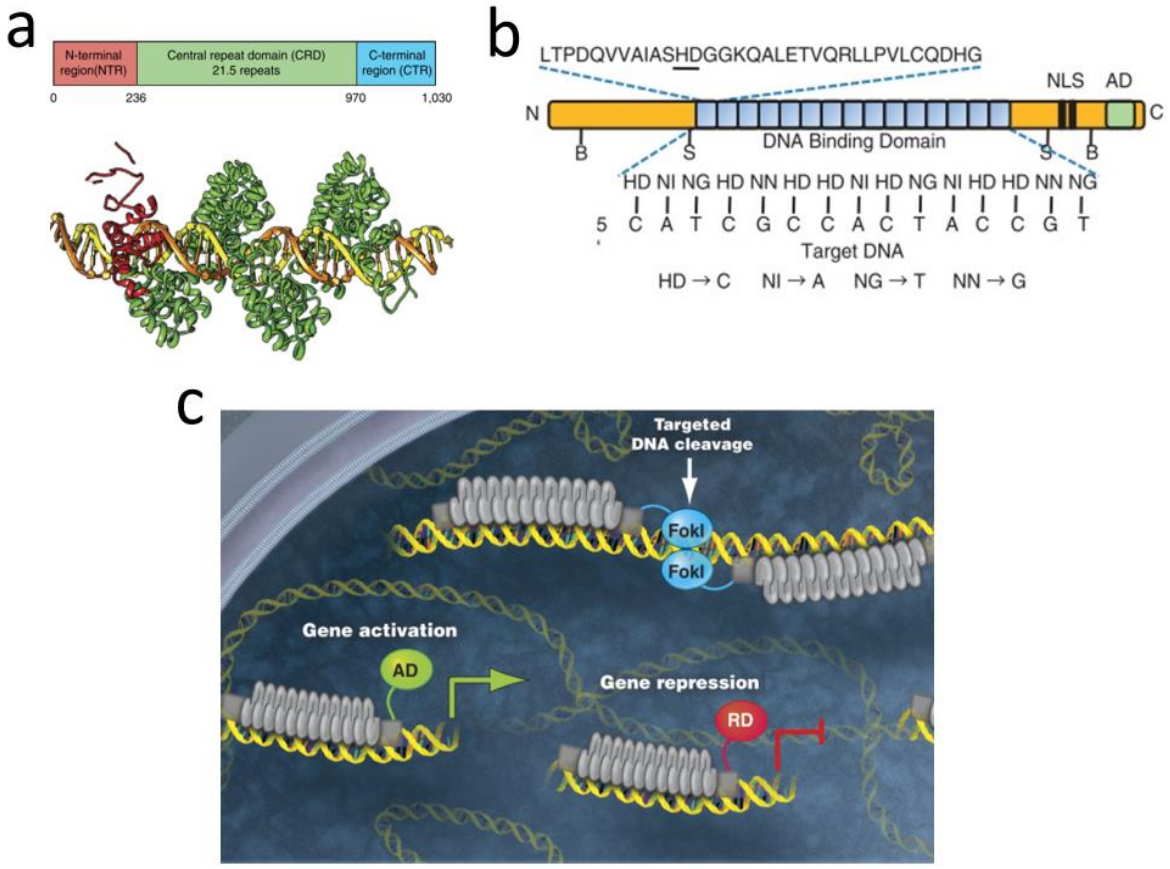


Figure 2.2. Transcription Activator-Like Effectors

(a) Schematic and ribbon diagram of a TALE protein binding to DNA. The protein wraps itself along the major groove (adapted from Cuculis *et al.* [1]. (b) Diagram showing the sequence and location of the DNA binding repeats with the repeat variable di-residue underlined. Amino acid binding code shown below (adapted from Cermak *et al* [2]. (c) Graphical representation of three ways designer TALEs can be used to influence DNA (adapted from Bogdanove *et al.* [6])

of the level of gene expression when paired with transcriptional activation or inhibitory domains, and the ability to identify and locate a gene within the nucleus *in vivo* when paired with an attached fluorophore (Figure 2.2c) [24-26]. dTALEs have been shown to be an attractive alternative to the CRISPR-Cas9 system when it comes to gene reactivation and activation, with evidence demonstrating that the dCas9 complex may impede the recruitment of factors in gene activation [27]. One of the features that makes the CRISPR-Cas9 system so attractive is the ease with which unique guide RNAs can be produced relative to the laborious process of designing a unique protein for each DNA target, as is required for dTALEs. Recently, a method of generating multiple dTALEs in a single day, without the need for serial bacterial transfections and selections, has been developed [28]. This method utilizes Golden-Gate cloning with a dsDNA library containing 60 base-determinant monomers, each encoding the amino acid domain and a unique RVD, flanked by universal primer annealing sites, and two linker domains. This binding region is then cloned into a sequence containing the TALE backbone and a CMV promoter for expression in mammalian cells.

It is known from N- and C-terminal deletion experiments performed on TALEs that the N-terminus contains an essential translocation domain while the C-terminal region is not necessary for dTALE-DNA interaction [29]. Single molecule experiments consisting of fluorescently labeled TALE proteins interacting with DNA extended and immobilized on a polyethylene glycol (PEG) surface have demonstrated that the search mechanism is not directionally biased and consists of two events [1]. The first event is micron scale hopping which involves

dissociation of the protein from DNA and the second event is slower 1-D diffusion along the DNA backbone [1]. This 1D diffusion can be broken down further into two binding modes: an initial non-specific rapid sliding search facilitated by the N-terminus and slower “recognition” events utilizing the DNA binding domain containing the sequence specific repeats. Further studies into the TALE protein search mechanism elucidated that the protein utilizes a rotationally decoupled mechanism during the fast 1D sliding search, enabling a robust interaction with DNA that still allows for rapid 1D diffusion [30]. These experiments were performed on random sequences of DNA that did not contain the target sequence for the TALE protein. Since there were no TALE target sequences present, these experiments were not able to characterize target sequence binding kinetics and the affinity and specificity of binding versus the number of repeats.

RESULTS

Photolithography as a DNA curtains patterning method

Previous DNA curtain flow cell fabrication methods have relied on electron-beam lithography to produce the patterns characteristic to DNA curtains, whether it be to create a mold for casting PDMS devices or nanoimprint lithography, or to directly deposit chrome on the glass surface. While electron-beam lithography provides sub 10 nm resolution necessary for many integrated circuit designs, it is not required to resolve the relatively large size (~200 nm) of DNA curtain rods. We selected photolithography as a tool to pattern devices as it is quicker, more cost effective, and the equipment is readily available to a larger set of the scientific community than electron-beam lithography.

Photolithography consists of several steps: first, a glass wafer is coated with photoresist, a material that becomes soluble in a particular chemical after exposure to deep ultraviolent (DUV) light (Figure 2.3a #1). Next, a photomask is created when a pattern designed on CAD software, in our case L-Edit, is printed onto a chrome and photoresist coated plate at 6 times the desired feature size (Figure 2.3b,c). After the photomask is exposed and processed, it is ready to be used to pattern features on the glass wafers. DUV illumination is directed through the photomask by an objective lens resulting in sub-micron features on the wafer below (Figure 2.3a #2). The exposed photoresist is then dissolved using a solvent, leaving bare glass to deposit chromium (Figure 2.3a #3). The partially photoresist-coated wafer is inserted into an evaporation chamber to coat it in a 20 nm layer of chrome, the desired final height of the nanobarriers (Figure 2.3a #4). A final liftoff step removes all remaining photoresist, leaving bare glass with chromium patterns (Figure 2.3a #5, d).

The feature size achievable by photolithography is limited by the wavelength of light, obeying the same principle as optical microscopy

$$d = k\frac{\lambda}{2NA} \tag{1}$$

Where k is a feature of the photoresist and the exposure process, λ is the wavelength of light, and NA is the numerical aperture of the lens. Photolithography tools commonly utilize deep ultraviolet light, λ = 248 nm, in order to resolve feature sizes on the order of 100 nm.

Our CAD design contained geometric wells with sharp edges in order to

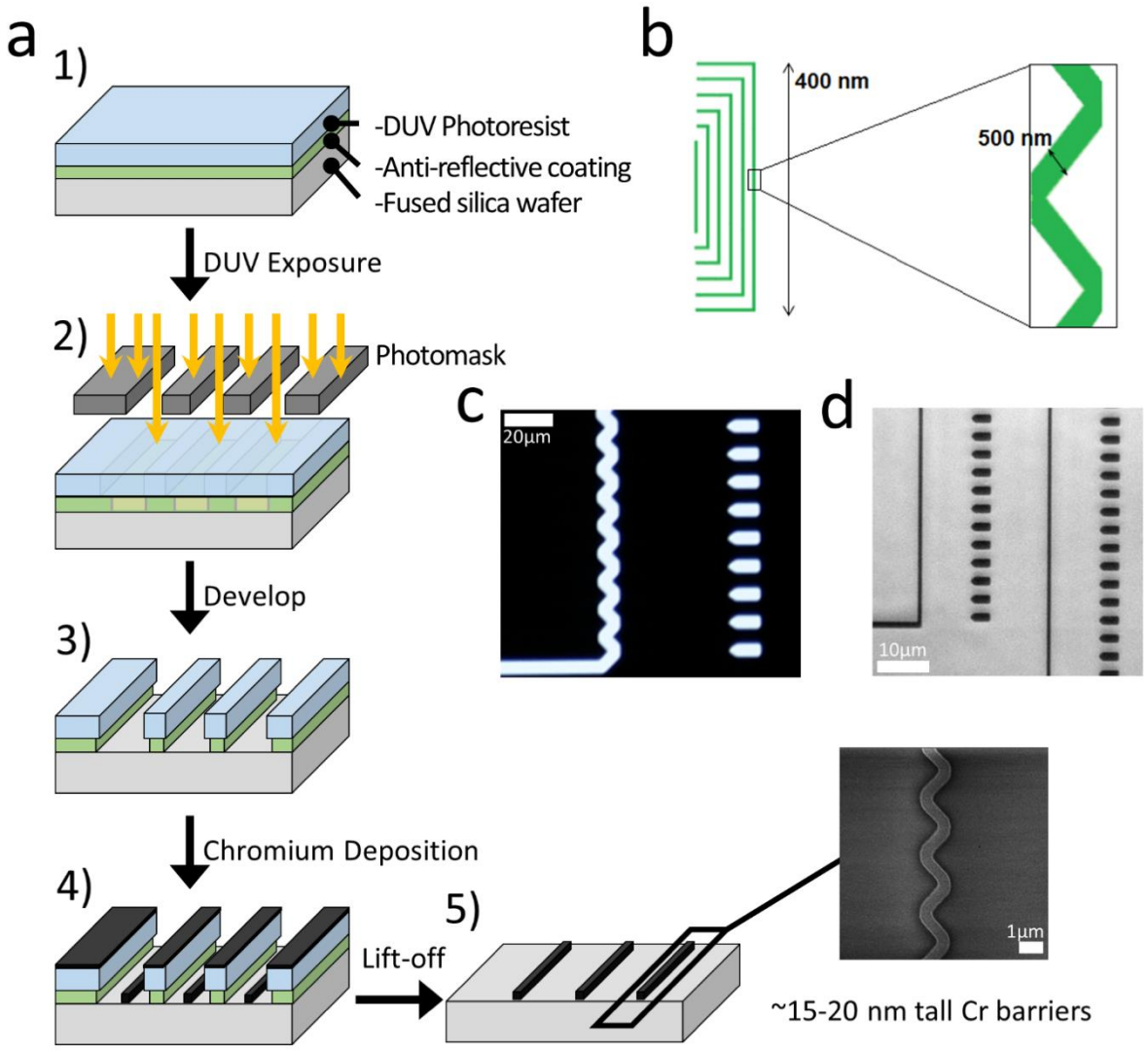


Figure 2.3. Nanofabrication of the DNA Curtain Nanopatterns
(a) Steps involved in photolithography. A photoresist-coated wafer is exposed to deep UV light through a photomask containing the desired patterns. Exposed photoresist is removed with developer, allowing chrome deposition on the glass. Liftoff removes all remaining photoresist leaving only chrome patterns on bare glass. (b) An example of one set of curtain rods in the design software, L-edit. Inset one nanowell blown up. (c) Darkfield image of the chrome photomask used in the exposure step (d) Brightfield image of the chrome patterns on glass prior to use in flow experiments.

separate individual DNA molecules in the “curtain” (Figure 2.3b). The photomask that resulted from this design lost resolution on the corners and produced rounded features which still resembled the nanowells (Figure 2.3c). After chrome deposition using a thermal/electron-beam evaporator and liftoff, we demonstrated that this fabrication method is sufficient to form sharply resolved 20 nm tall DNA curtain rods down to a width of 500 nm (Figure 2.3d, 3a). The nanopillars used as the anchors for double tethering were scaled up in size from the original 500 x 370 nm^2 to 2000 x 1000 nm^2 in order to offset the increased surface area of the curtain rods (Figure 2.3c,d).

The limitations of photolithography are demonstrated by the rounded, sinusoidal pattern seen in place of the sharp peaks and troughs of the sawtooth pattern designed to create spacing between DNA molecules (Figure 2.3c,d, 2.4a). However, we will show that the sinusoidal pattern we produced is sufficient to form regularly spaced DNA curtains.

Bilayer composition and fluidity

Bilayers composed primarily of 1,2-dioleyl-sn-glycero-3-phosphocholine (DOPC) containing a small fraction of biotin labeled 1,2-dioleyl-sn-glycero-3-phosphoethanolamine (DOPE) have been shown to prevent nonspecific adsorption of biomolecules to glass while allowing lipids to diffuse freely along the surface. The addition of PEGylated DOPE provides another barrier to nonspecific surface adsorption. We sought to find an ideal PEG chain length and lipid fraction to maximize the surface passivation effect without compromising bilayer fluidity. In

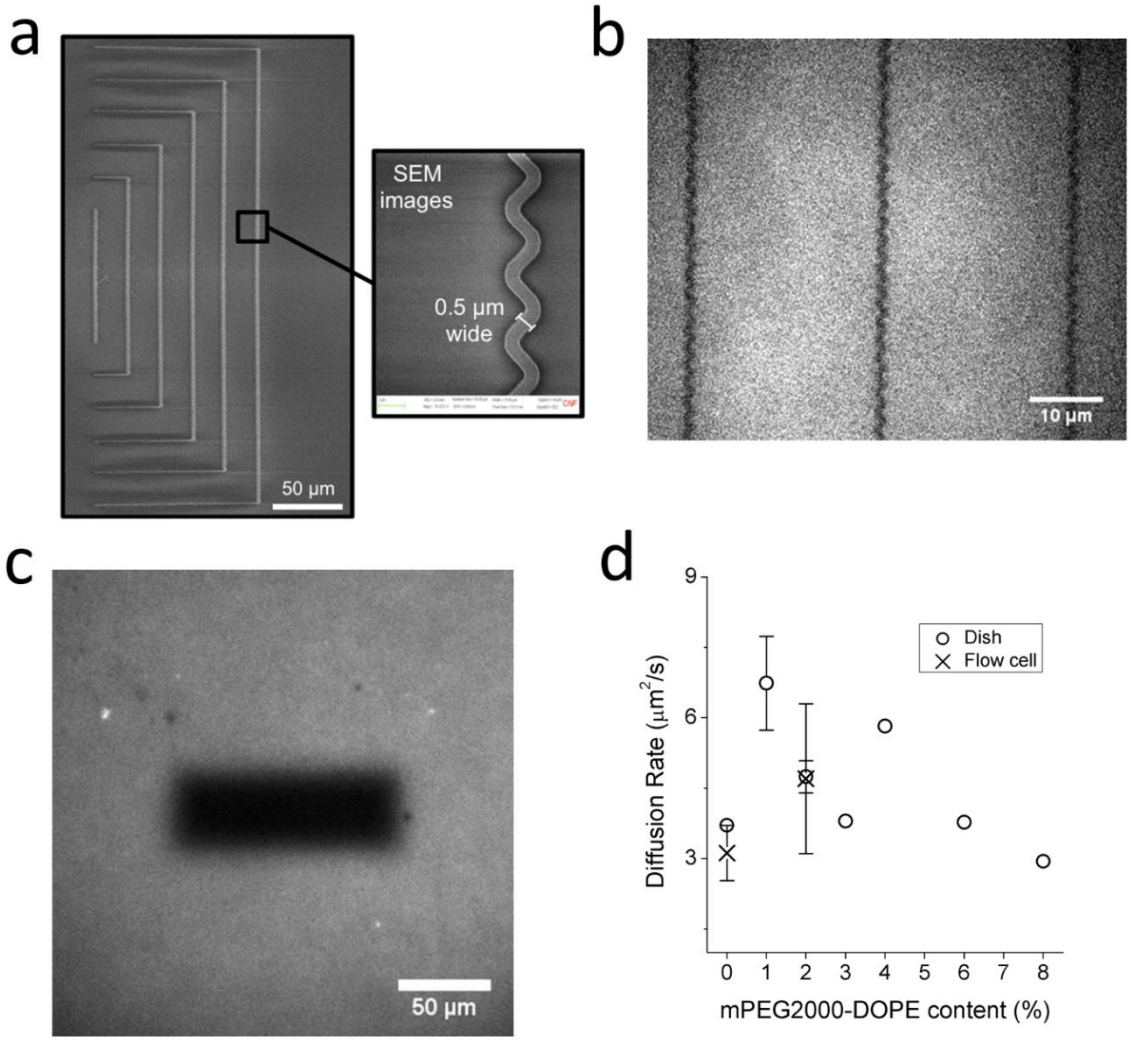


Figure 2.4. Flow Cell Patterns and Bilayer Composition
(a) SEM image of the arrays of curtain rods with an expanded view of the rounded edges of the sawtooth pattern produced by photolithography. (b) Fluorescent lipid bilayer containing 0.1% Lissamine Rhodamine-DOPE on the surface of the flow cell between curtain patterns (c) Example of a photobleached rectangle in a bilayer formed on a MatTek dish to measure diffusion rates. (d) Diffusion times for different amounts of mPEG550-DOPE both in the flow cell and in MatTek dishes. n=3 for 0,1, and 2%. n=1 for others.

order to visualize the bilayer, 0.1% Lissamine Rhodamine-DOPE was added to each lipid mixture. Surface preparation methods and bilayer deposition conditions were also tested to determine the best protocol to form fluid lipid bilayers.

A clean and hydrophilic surface is required for the formation of a lipid bilayer as surface irregularities can introduce gaps in the bilayer and the hydrophilic headgroups of lipids on the "inner" surface of the bilayer need to interact strongly with the surface [7]. A 15 minute glass cleaning with 1M NaOH and 3 minutes in oxygen plasma both produced surfaces necessary for a continuous and fluid bilayer to form in a robust fashion on a glass bottom dish and inside the DNA curtains flow cell (Figure 2.4b,c). Since the DNA curtains devices are reusable, plasma cleaning is the preferred method as repeated exposure to NaOH may dissolve the glass surrounding the chromium patterns and cause them to detach. Multiple lipid incubation steps are included in our flow cell preparation protocol to ensure every part of the surface has a bilayer present since any gaps will prevent a barrier to lipid diffusion and introduce a point for continued buffer flow to further disrupt the bilayer. This process may result in areas with stacked lipid bilayers, but that does not interrupt the ability for lipids to diffuse. A vigorous rinse with either lipid buffer or MilliQ H_2O was required to remove unfused liposomes from the surface of the bilayer, although care must be taken to prevent oxygen exposure that irreversibly damages the bilayer.

The concentration of PEGylated lipids is limited only by bilayer fluidity as the large PEG molecules form an increasingly dense network of hydrophilic chains across the bilayer. mPEG550 and mPEG2000 were both tested at concentrations

up to 8% mole fraction. To evaluate bilayer mobility as a function of PEGylated lipid concentration and ensure the flow cell surface quality was sufficient to promote bilayer formation, we used fluorescence recovery after photobleaching (FRAP) measurements. FRAP measurements were taken of each lipid mixture, with 0.1% Lissamine Rhodamine-DOPE added for visualization, deposited either on a MatTek dish or the inner surface of the flow cell to determine the rate of diffusion (Figure 2.4c). The fluorescence intensity recovery for rectangular photobleaching patterns is given by the equation

$$I(t) = I_f(1 - \sqrt{w^2(w^2 + 4\pi Dt)^{-1}}) \tag{2}$$

where t is the time in seconds after photobleaching, $I(t)$ is the intensity at time t, I_f is the maximum recovered fluorescence, w is the strip width, and D is the diffusion rate [31, 32]. The time required for 50% of the fluorescence to recover, $\tau_{1/2}$, is a measured quantity described by

$$I(\tau_{1/2}) = \frac{I_f}{2} \tag{3}$$

We can solve for diffusion time by setting $t = \tau_{1/2}$ in equation (2)

$$I(\tau_{1/2}) = I_f(1 - \sqrt{w^2\left(w^2 + 4\pi D\tau_{1/2}\right)^{-1}}) \tag{4}$$

$$\frac{I(\tau_{1/2})}{I_f} = (1 - \sqrt{w^2\left(w^2 + 4\pi D\tau_{1/2}\right)^{-1}}) \tag{5}$$

With some rearranging of terms we get

$$D = \frac{w^2}{4\pi\tau_{1/2}}\left[\left(1 - \frac{I(\tau_{1/2})}{I_f}\right)^{-2} - 1\right] \quad (6)$$

because $I(\tau_{1/2}) = \frac{I_f}{2}$, equation (6) becomes

$$D = \frac{w^2}{4\pi\tau_{1/2}}\left[\left(1 - \frac{1}{2}\right)^{-2} - 1\right] = \frac{3w^2}{4\pi\tau_{1/2}} \quad (7)$$

The rate was found to be ~5 $\mu m^2 s^{-1}$, indicating that the lipids were very mobile. Minimal change in diffusion time was observed between the two lengths of PEG chains and mole fractions ranging from 0-8% (Figure 2.4d). Furthermore, there was no significant change observed in diffusion time between bilayers formed in a MatTek dish versus those formed in the DNA curtains flow cell (Figure 2.4d). We decided to go with 3% mPEG550-DOPE for our protein binding experiments.

Preparation of the DNA curtains flow cell

With verification that the passivating bilayer was intact and the lipids were freely diffusing throughout, the next step was to ensure biotinylated lambda DNA bound to the bilayer was able to diffuse in the presence of buffer flow to the nanobarriers in order to form uniformly extended DNA curtains (Figure 2.5a). To verify that the DNA was tethered only at the end and not adsorbed to the surface along the length of the molecule, we performed single tether experiments with periodic application of hydrodynamic force (Figure 2.5b). 8 pM lambda DNA was introduced to the chamber and allowed to bind to the bilayer in the absence of flow (Figure 2.5b upper panel). The hydrodynamic force was then applied by flowing

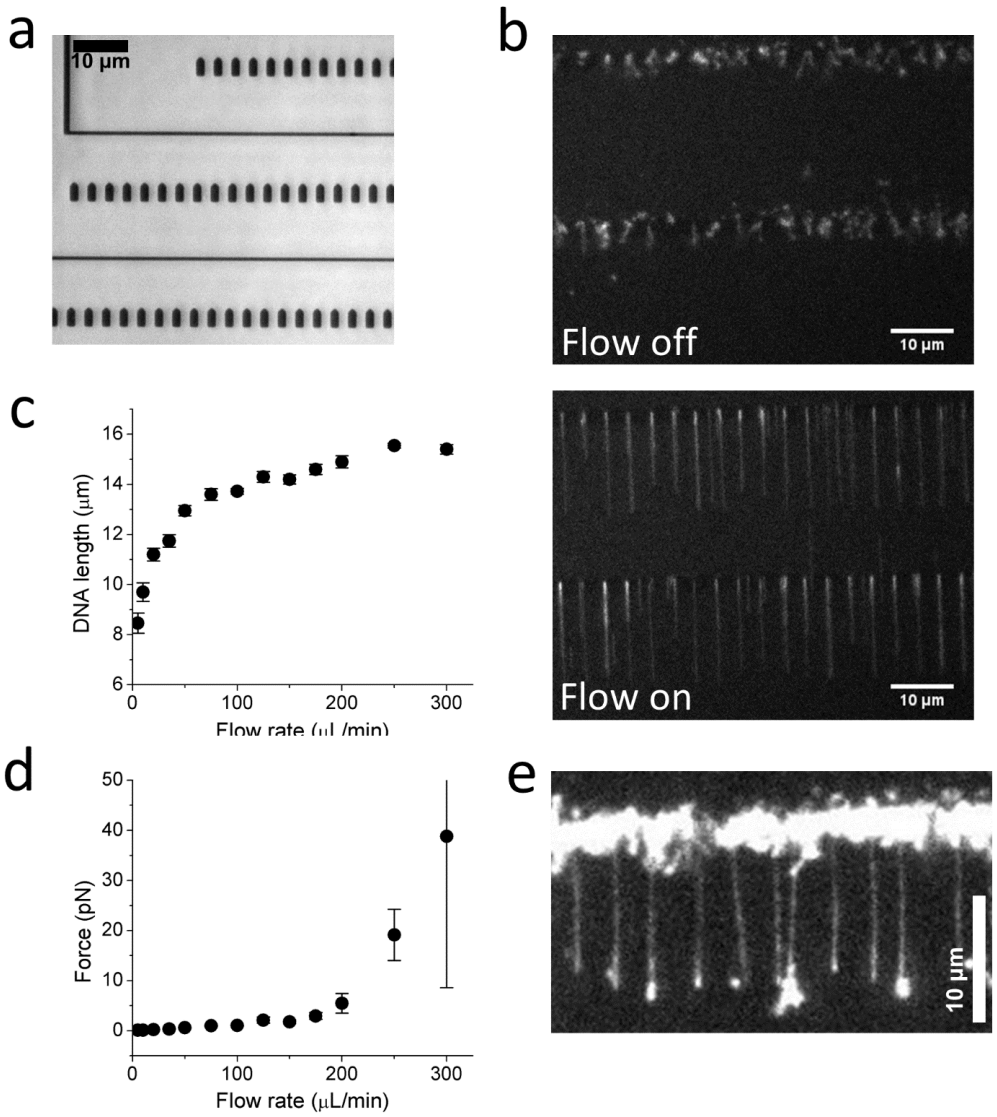


Figure 2.5. Behavior of DNA in the Flow Cell
(a) Brightfield image of the nanobarriers and nanopillars of a double tether curtain pattern. Curtain rod positioned 13 µm above the nanopillars. (b) Demonstration of the DNA freely diffusing up out of the TIRF field in the absence of flow (upper panel) and then extended and forming a curtain when flow is resumed (lower panel). (c) A plot of DNA extension vs buffer flow rate (n=5). (d) Converting extension into force using the WLC model, a plot of force vs buffer flow rate (n=5). (e) Double tethered DNA molecules, stained with YOYO-1, extended in the absence of flow.

imaging buffer containing YOYO-1 through the flow cell using a syringe pump and the DNA molecules extended in the imaging plane as expected (Figure 2.5b lower panel). This was repeated several times, and each time the DNA freely diffused upwards out of the plane of illumination in the absence of force, and then extended parallel to the surface when force was applied. Furthermore, we observed no fluorescent signal apart from the DNA molecules, which implies that the surface is sufficiently passivated to prevent nonspecific adsorption of biomolecules or impurities. The force applied to DNA molecules in microfluidic devices is not as precise as those which can be achieved with force spectroscopy techniques such as optical and magnetic tweezers. In this experiment, we sought to quantify the approximate forces exerted as a function of flow rate. First, we measured the length of the DNA molecule at a given flow rate to quantify the extension (Figure 2.5c). By measuring the DNA extension and fitting it to the worm-like chain model

$$\frac{FL_c}{k_BT} = \frac{1}{4}\left[\left(1 - \frac{x}{L_o}\right)^{-2} - 1\right] + \frac{x}{L_o} - 0.8\left(\frac{x}{L_o}\right)^{2.15} \quad (8)$$

where F is the force applied, L_c is the characteristic length (42.5nm for DNA), k_B is Boltzmann's constant, T is the temperature, L_o is the contour length of the DNA molecule (16.2 μm for lambda DNA), and x is the linear distance between the two ends of DNA, the force applied at a given flow rate could be calculated [33, 34]. The resultant force was determined to range from 0.1 – 15 pN which is on the order of forces experienced by DNA *in vivo* (Figure 2.5d) [35]. We decided to use 100-200 μL/min (1-5pN) for our DNA extension experiments as they occur on the flow-force plateau and extend DNA to 80-90% of its contour length while causing minimal DNA breakage or disassociation from the bilayer.

If we planned to perform a double tether DNA curtains experiment, anti-digoxigenin fab fragments were incubated in the flow cell to nonspecifically cover all of the chrome patterns. Given the relative size of the double tethering anchors compared to the curtain rods themselves, minimal “looping” of DNA, where both ends were interacting with the curtain rod, was observed. The DNA molecules remained extended in the absence of force, signaling successful double tethering (Figure 2.5e).

Proteins bound to DNA curtains

Successful DNA-protein binding in the flow cell without nonspecific surface adsorption was assessed using two well characterized proteins, EcoRI^{E111Q}-3xFLAG and LacI-3xFLAG. We used the catalytically inactive variant of EcoRI^{E111Q} so that binding could be visualized without nuclease activity. Wild type lambda DNA contains 6 EcoRI binding sites sufficiently separated from one another so that at least 4 could be individually identified.

The ability to insert desired sequences into lambda DNA expands the flexibility of this powerful technique. In order to evaluate whether we could successfully insert non-native protein binding sites into lambda DNA, and have proteins bind to them, we inserted two repeats of the *LacO* sequence separated by a short spacer 20 kilobases away from the biotinylated end of lambda DNA.

The function of both LacI-3xFLAG and EcoRI^{E111Q}-3xFLAG was confirmed via electrophoretic mobility shift assay (EMSA) with an approximate respective k_D of 200 nM and 5 nM (Figure 2.6 a,c). These values are significantly higher than published data, but this could potentially be due to the presence of the 3xFLAG

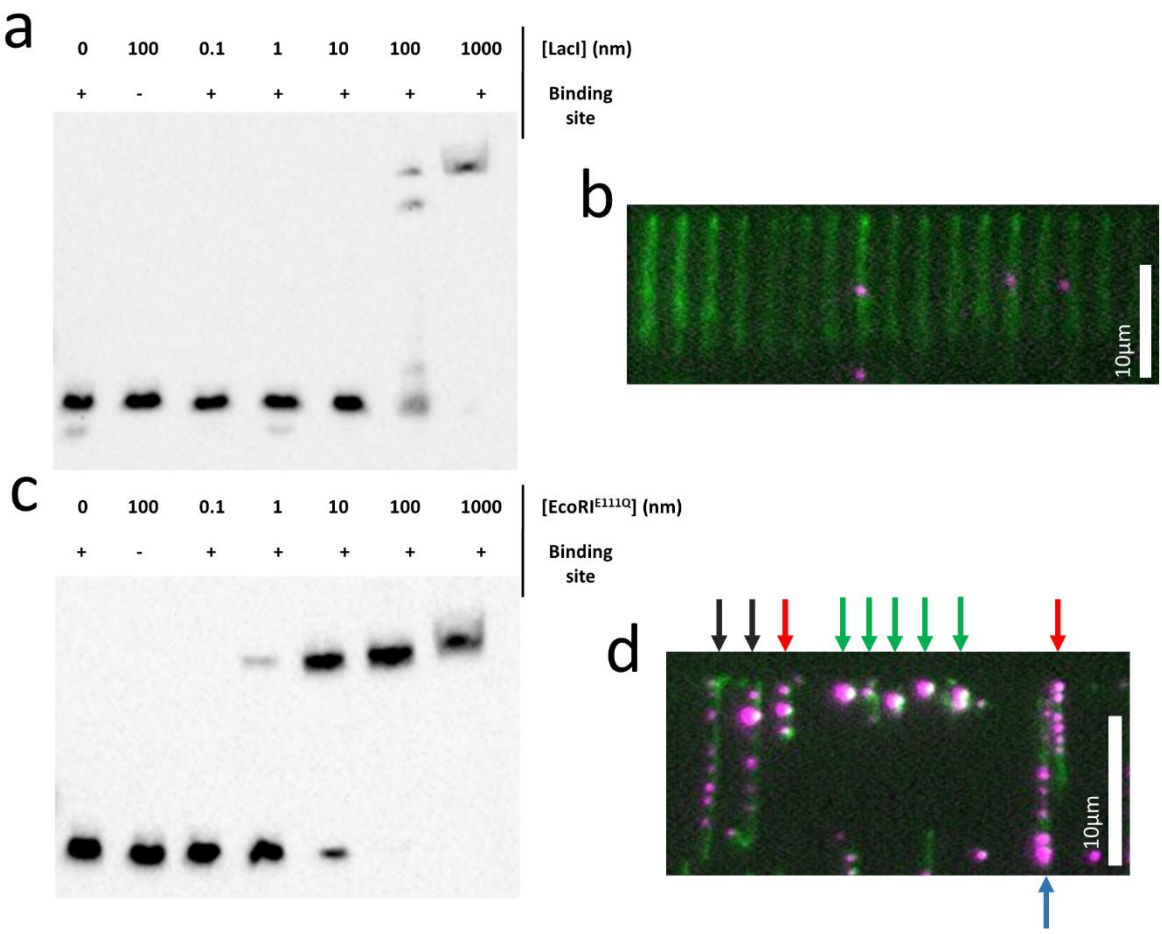


Figure 2.6. Protein Binding to DNA Curtains

(a) PAGE gel result of an EMSA to measure the approximate k_D of LacI-3xFLAG using 20 fmol of biotinylated DNA labeled with a streptavidin-horseradish peroxidase conjugate. (b) DNA curtain (stained with YOYO-1) showing bound LacI-3xFLAG labeled with Alexa 647 (c) PAGE gel result of an EMSA to measure the approximate k_D of EcoRI^{E111Q}-3xFLAG (d) Double tethered DNA curtain showing bound EcoRI^{E111Q}-FLAG labeled with Alexa 647. Black arrows indicate DNA molecules with expected binding. Red arrows indicate "looped" DNA. Blue arrow indicates protein building up at the downstream end of the DNA molecule. Green arrows indicate molecules which did not double tether, and diffused upward in the absence of flow.

tag or the composition of binding buffer [14, 36]. Furthermore, the proteins did not bind to a control DNA strand lacking a binding site. This evidence is sufficient to demonstrate that these are adequate test proteins to use in the flow cell.

LacI-3xFLAG labeled with an Alexa 647 conjugated FLAG antibody was incubated with DNA prior to injection into the flow cell. It appears that an insufficient amount of protein was used, since very few DNA binding sites were occupied by protein (Figure 2.6b). Bound protein may have been caused to dissociate due to force applied during the injection and washing steps necessary to remove unbound molecules. However, the proteins that remained bound were in the correct location on the DNA molecule, which indicates that foreign DNA sequences were successfully inserted into lambda DNA and that proteins can bind their target inside the conditions present in the flow cell.

The experiment was repeated using EcoRI^{E111Q}-3xFLAG on double tethered curtains. Our observations suggest that too much protein may have been used, as some binding sites were occupied by multiple proteins (Figure 2.6d). In a double tethered setup, some fraction of DNA molecules don't successfully tether their second end, and a buildup of protein on untethered DNA may explain the large magenta and green puncta in the image (Figure 2.6d green arrows). To support the argument of flow interfering with protein interactions, we found that EcoRI proteins were pushed to the end of the DNA molecule, producing large magenta puncta (Figure 2.6d blue arrow). It was difficult to determine whether the binding sites were specifically occupied due to the large amount of protein in the flow cell, but some of the proteins that did bind were located in the correct regions

of the DNA molecule (Figure 2.6d black arrow). As with the LacI construct, the flow applied after binding may have caused proteins to dissociate. This supports the use of double tether curtains as a method to view protein dynamics in the absence of external force.

Transcription Activator-Like Effectors

TALE proteins have the ability to bind any sequence of DNA depending on the identity of the RVD present in each of the 34 amino acid repeats that form the DNA binding domain (Figure 2.7a). TALE1535, designed by Fabio Cupri Rinaldi of the Bogdanove laboratory at Cornell University, has been shown to bind its target with a picomolar affinity, thereby providing a very tightly binding specimen with which to elucidate the search mechanism of TALE proteins (Figure 2.7c). The TALE1535 binding site was inserted into lambda DNA at the same location and using the same method as the *LacO* sites above (Figure 2.7b). TALE protein was incubated with DNA prior to injection into the flow cell in an attempt to maximize binding to the target sequence. The protein-DNA mixture was injected into the flow cell at flow rates ranging from 50-400 μL/min. Protein concentrations ranging from 30-1500 nM (~1-100 nM labeled protein) were tested, with nonspecific binding to the chrome barriers proving to be a significant challenge at all tested concentrations. To further address this, multiple buffers with various ionic strengths and additives to reduce nonspecific interactions were tested; however, the adsorption remained an issue. Flow experiments with the previously incubated protein-DNA complexes revealed that TALE protein bound specifically to the target region with some positional variation likely due to non-uniform DNA extension

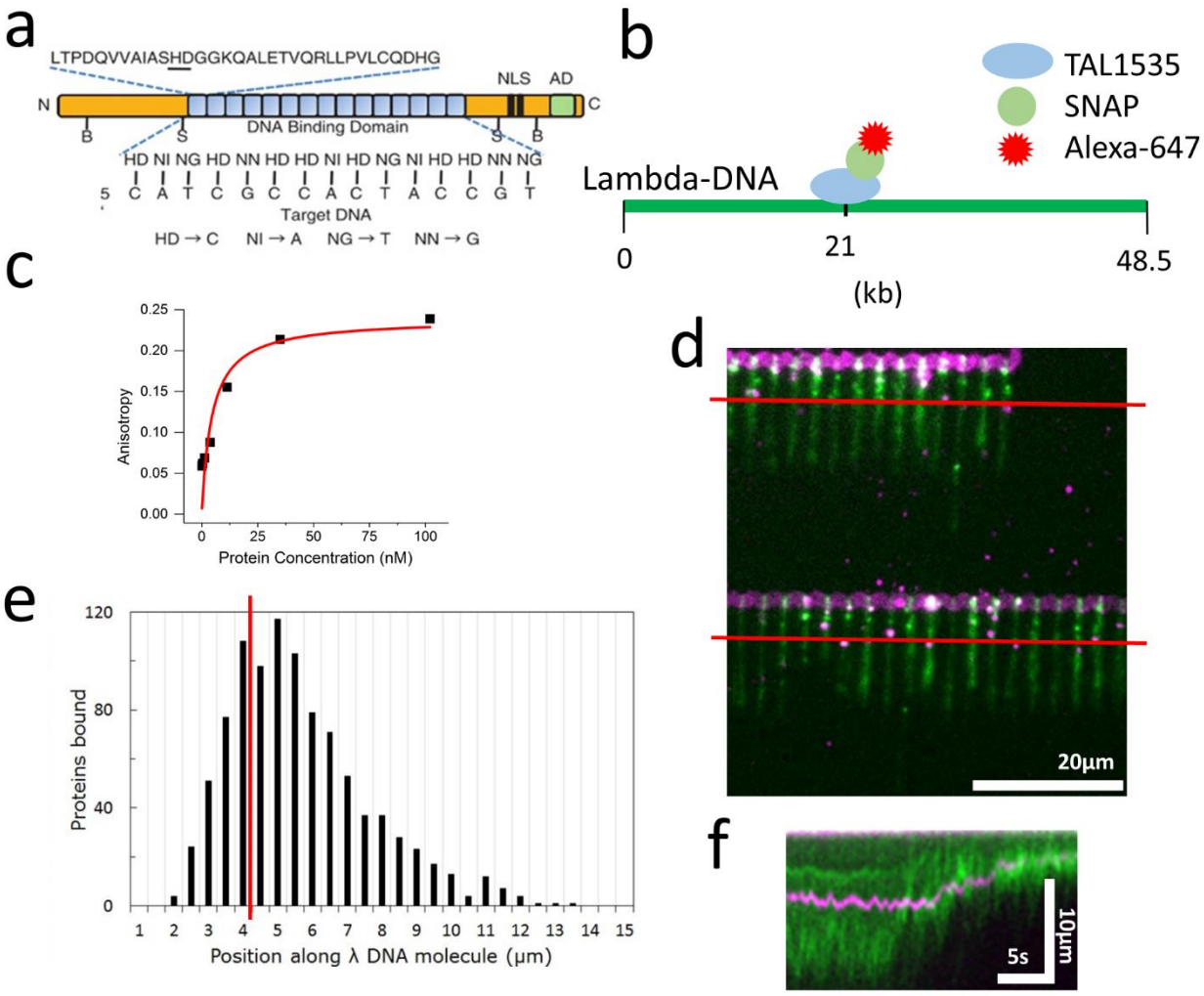


Figure 2.7. Transcription Activator-Like Effector Binding Experiments
(a) Example of a TAL protein showing the 34 amino acid repeats that form the DNA recognition region, with the RVD underlined. Common amino acid sequences used to bind each DNA base shown below. (b) Schematic of labeled TAL1535-SNAP binding to TAL1535-lambda showing the location of its binding site along the linear DNA molecule. (c) Fluorescence anisotropy data to determine the k_D of TAL1535-SNAP. (d) Image of Alexa 647 labeled TAL1535-SNAP binding to YOYO-1 stained DNA curtains. Red line overlay corresponds to expected binding location assuming 90% extension of DNA.(e) Histogram of the measured distances from the biotinylated end of DNA to bound protein (n~400) Red line overlay corresponds to expected binding location assuming 90% extension of DNA. f) Kymograph showing the stability of the bound TAL protein as flow was relaxed and DNA diffused out of the field of illumination

(Figure 2.7d,e). We posit that any protein which interacted with DNA was bound tightly enough to remain immobile, while proteins that may have been weakly interacting were washed off and adsorbed to the chromium nanobarriers, thus no dynamics were observed (Figure 2.7f).

A second experiment was carried out using a mNeonGreen-TALE1456 designed to bind a site in native lambda. Data from this experiment is presently unavailable, but these results also showed a large fraction of the protein adsorbed to the chrome rods as with TALE1535. From these experiments, we determined that TALE proteins were not compatible with the DNA curtains technique due to significant nonspecific interactions between the protein and the chromium patterns.

CONCLUSION

We demonstrated that the nanobarriers and polygonal anchors that constitute DNA curtains single- and double-tethered devices can be patterned using photolithography at a resolution sufficient to create an array of wells that immobilize and separate individual DNA molecules. This serves to make the technique available to a wider range of experimentalists.

We also demonstrated that a primarily DOPC bilayer formed in the flow cell can have an 8% mole fraction of mPEG550-DOPE or mPEG2000-DOPE without losing fluidity, enabling a range of concentrations to be used depending on the conditions which best suit the protein of interest.

Finally, we demonstrated that protein binding within our flow cell on double tethered DNA was achievable, using LacI, EcoRI^{E111Q}, TALE1456, and TALE1535.

It is my conclusion that the high affinity of TALE proteins for the chrome curtain rods disqualifies them from the list of potential proteins that can be visualized in the DNA curtains flow cell.

METHODS

The work pertaining to nanofabrication, flow cell assembly, liposome preparation, and initial DNA curtain formation/flow-extension measurements was performed in collaboration with Devin Wakefield of the Baird-Holowka Laboratory at Cornell University.

Nanobarrier design and fabrication

All of the steps from pattern design to photoresist lift-off utilized equipment made available to us in the Cornell Nanoscale Science and Technology Facility (CNF).

A photomask containing the single and double tether curtain patterns in 0.5, 1, and 1.5 μm feature thicknesses was designed in L-Edit and written onto a photosensitive chromium mask plate (plate provided by CNF) using the CNF DWL 2000 High Resolution Pattern Generator (Heidelberg Instruments). Exposed photoresist-coated chromium was stripped using the CNF Mask Developer Tool. The substrate to be patterned was a 100 mm diameter and 1 mm thick fused silica wafer (Sydor Optics). This was first coated with a 40 nm thick layer of anti-reflective coating (DS-K101-4 (Brewer Science)) followed by a 300 nm thick layer of photoresist (UV210 (Dow Chemical Co.)), both using the CNF Gamma Cluster Tool (SUSS MicroTec). The patterns were transferred to the photoresist coated wafer via exposure to 248 nm light through the previously designed photomask

using the CNF ASML PAS 5500/300C DUV Stepper. Exposed photoresist was removed using a photoresist developer (AZ 726 MIF) utilizing the CNF Gamma Cluster Tool exposing the glass substrate for chromium deposition. A 20 nm thick layer of chromium was deposited using the CNF SC4500 Thermal/Electron-beam evaporator (CVC) and the wafer was diced into three 60 x 25 mm^2 rectangles, each containing 16 sets of patterns, using the CNF dicing saw (K&S 7100). Finally, the unexposed photoresist was removed via an overnight lift-off step using Microposit Remover 1165, leaving a glass slide containing patterns in the central region that is suitable for flow cell formation and imaging.

Flow cell assembly

The patterned glass substrate of the DNA curtains flow cell was prepared for use by drilling inlet and outlet holes with a 1.5 mm diamond tipped drill bit (Kingsley North) using a rotary tool (Dremel®). After cleaning the glass surfaces with Hellmanex (Hellma GmbH & Co.), 1M NaOH, 100% ethanol, MilliQ H_2O and drying with nitrogen, the slides were further cleaned with oxygen plasma (Harrick Plasma Basic Plasma Cleaner) for 5 minutes along with coverslips. Double sided tape (3M) with a rectangular cutout of the flow cell profile was sandwiched between the patterned glass and a coverslip. This was heated to 75°C in an oven for 30 minutes to promote bonding. NanoPorts (IDEX Health & Science) were attached using hot glue (Surebonder hot glue gun) and heated on a hot plate to allow the glue to flow and form a water-tight seal.

Liposome preparation

Liposomes were made from 95% DOPC, 4.4% mPEG2000-DOPE, 0.6% biotin cap-DOPE with 0.1% Lissamine Rhodamine-DOPE added for fluorescent bilayers (given in mole percentages) (Avanti Polar Lipids, Inc). A mixture of the above lipids was made from chloroform stocks and dehydrated under nitrogen until visibly dry, ~3 minutes, before further dehydration in a vacuum (Harrick Plasma Basic Plasma Cleaner) overnight. This was rehydrated with lipid buffer (50 mM Tris pH 8.0, 100 mM NaCl) and allowed to rehydrate on ice for 10 minutes. The mixture was vortexed for 1 minute partway through this rehydration period. The aqueous lipid solution was then sonicated using a probe sonicator (Misonix Inc.) for 20 minutes on ice, centrifuged at 16000xg for 5 minutes and filtered to remove any debris from sonication.

For initial bilayer composition and fluidity tests, liposomes were deposited on glass coverslip bottom dishes (MatTek Corporation) after plasma cleaning and allowed to coat the surface for 20 minutes before gentle but thorough rinsing with MilliQ H_2O. Bilayers were formed in the DNA curtains flow cell as described below.

Bilayer diffusion imaging

Bilayer visualization and FRAP measurements were performed on an Andor revolution spinning disk confocal microscope that is part of the Cornell Biotechnology Resource Center (NIH S10OD010605). An Olympus UPLSAPO 10x/0.4 Air objective lens was used for these experiments. The excitation and photobleaching of Lissamine Rhodamine was done with 561 nm illumination. A rectangular region of interest (ROI) was selected and photobleaching was

achieved by exposure to maximum laser intensity for ~15 ms. Images were taken every 10 seconds afterwards for 3 minutes in order to observe the recovery of fluorescence in the depleted region.

End labeling of DNA for single molecule experiments

In order to attach the DNA to the bilayer, the right end of linearized lambda phage DNA (New England Biolabs, Inc) was hybridized with a biotinylated 12 bp DNA oligomer: GGGCGGCGACCT-BioTEG (Integrated DNA Technologies, Inc). The left end of lambda was left unlabeled for single tethered experiments, but hybridized with a digoxigenin labeled 12bp oligo: AGGTCGCCGCCC-DigN (Integrated DNA Technologies, Inc) for double tether experiments. Oligomer annealing was carried out by combining 3 nM lambda DNA and 200 nM oligomer in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA), heating to 80°C for 5 minutes and gradually cooling to room temperature. T4 DNA ligase and ligase buffer (New England Biolabs, Inc) were added and incubated at 16°C overnight. If both ends were going to be labeled, then this process was repeated using 360 pM single end labeled lambda DNA and 36 nM oligomer.

Preparation of flow cell with lipid bilayer and DNA

DNA curtains flow cells were prepared as previously described, with some adaptations [13]. The assembled chamber was first flushed with lipid buffer (10 mM Tris pH 7.8, 100 mM NaCl), then 1 mL of prepared liposomes at 1 mg/mL in lipid buffer were injected in 3 stages, with 5 minutes between injections. The chamber was flushed with lipid buffer to remove unbound liposomes and the remaining liposomes were allowed to settle and spread on glass for 20-30 minutes.

If double tethering was going to be performed, 1 mL of 80 µg/mL Anti-Digoxigenin Fab fragments (Sigma-Aldrich) in PBS was injected in three stages with10 minutes between injections. To block surfaces from nonspecific binding, the chamber was flushed with BSA buffer (40 mM Tris pH 7.8, 1 mM $MgCl_2$, 1 mM DTT, 0.2 mg/mL BSA, filtered and degassed under vacuum) and then allowed to rest for 20 minutes. 1 mL of lambda DNA at ~10 pM in 1 mL of BSA buffer (~10 fmols) was incubated over 3 stages with 10 minutes between to allow the biotin-streptavidin linkage to occur. Prior to imaging the chamber was flushed with imaging buffer (BSA buffer + 20 µL GLOX (Sigma Aldrich), 200 µL 40% Glucose) using a syringe pump (Harvard Apparatus) to remove any unbound DNA.

Imaging of DNA and proteins in the flow cell

Hydrodynamic force was generated via continuous flow of imaging buffer (with 5 nM YOYO-1 (Thermo Fisher Scientific) added if DNA was to be visualized) provided by a syringe pump (Harvard Apparatus). The DNA was moved to “curtain rods” and extended by continuous buffer flow for 5-10 minutes to allow the anti-digoxigenin to bind its hapten and anchor the second end of DNA. Depending on the experiment, flow was either continuously applied, or stopped during imaging.

Imaging of DNA and proteins within the flow cell was done on a custom microscope described in further detail in Appendix 1. The objective lens used was an Olympus UPlanApo 60x/1.2 Water. The excitation wavelengths used were 488nm for YOYO-1, 561 nm for Lissamine Rhodamine-DOPE, and 640 nm for Alexa-647. The filter cube (Andor) used to split fluorophore emission on to two cameras contained a 580 nm dichroic with 525/40 nm filter to collect YOYO-1

emission and a 679/41 nm (filters from Chroma Technology) filter to collect Alexa-647 emission. The beam was rotated at a low angle of deflection to ensure uniform illumination of the field and average out any optical aberrations.

Inserting binding sequences into lambda DNA

In order to introduce designer DNA sequences into lambda DNA, a 4 kb segment was removed by restriction digest with NgoMIV and XbaI followed by gel purification. A dsDNA molecule containing either 2x LacO binding sites (AATTGTTATCCGCTCACAATT) with a 10 bp spacer or the binding site for TAL1535 (CTGGATCATTCCCGAGCGCT) and with 2 bp overhangs on each end matching the digestion product of AgeI and NheI was designed and formed by annealing DNA oligomers (IDT). One of these DNA fragments was added to the two large lambda fragments and ligated to form a 44.5 kb designer lambda molecule titled 2xLacO-lambda TAL1535-lambda respectively.

Selection and amplification of successfully ligated DNA was performed by using MaxPlax Lambda Packaging Extracts (Lucigen) to create active phage molecules. These were used to infect the LE392 strain of *E. coli* grown on NZCYM broth (Sigma-Aldrich) agar plates to induce lysis upon infection so that plaques could be identified and selected. Large scale harvesting of DNA was performed by inoculating a suspension culture of LE392 in NZCYM with a plaque and allowing it to grow at 37°C while shaking. The OD600 was monitored and after it rose above 1.0 and returned to ~0.3, chloroform was added at 0.02% v/v. The culture was collected; NaCl was added to a final concentration of 1M and incubated on ice for 10 minutes. Centrifugation at 12000 xg for 10 minutes was performed twice

sequentially, collecting the supernatant each time. In order to precipitate out the phage particles PEG8000 was added at 10% w/v and incubated on ice for 30 minutes. Two sequential centrifugation steps at 12000 xg for 15 minutes each were performed, discarding the supernatant each time. The pellet was resuspended in PD buffer (10 mM Tris-HCl pH 7.6, 100 mM NaCl, 10 mM $MgCl_2$). RNase A (Sigma Aldrich) and DNase (Promega Co.) were added to 20 µg/mL and 5 µg/mL respectively final concentrations and incubated at 37°C for 30 minutes to remove any large fragments of nucleic acid not contained within a capsid. In order to break down the phage capsid, one volume of 0.3 M Tris-HCl plus EDTA to 100 mM, SDS to 1.25%, and proteinase K (New England Biolabs, Inc.) to 50 µg/mL and incubating for 10 minutes at 65°C. One third volume of 3M potassium acetate was added and incubated on ice for 10 minutes to precipitate proteins from solution. This was then centrifuged at 8000 xg for 10 minutes, the supernatant was collected, and 0.7 volumes of isopropanol were added to precipitate the DNA. This was centrifuged at 8000 xg for 10 more minutes and the supernatant was discarded. The DNA pellet was rinsed with 70% ethanol and resuspended in TE buffer. Phenol-chloroform extraction was performed by adding one volume of 25:24:1 phenol:chloroform:isoamyl alcohol, shaking by hand, and centrifuging at 16000xg for 5 min. The aqueous phase was collected and the process was repeated once more. Phenol was removed from solution by performing the same steps as above three times using chloroform. Finally, DNA was precipitated out using one volume of isopropanol, rinsed with 70% ethanol, and resuspended in TE buffer.

These protocols were provided by Daniel Duzdevich of the Greene Laboratory at Columbia University.

Protein expression

TALE1535-SNAP was cloned, expressed, and purified by Fabio Cupri Rinaldi of the Bogdanove Laboratory at Cornell University.

Plasmids for EcoRI^{E111Q}-3xFLAG and LacI-3xFLAG with a chitin-binding domain affinity tag were generously provided by Sy Redding of the Greene Lab at Columbia University.

Cells from the BL21 strain of *E. coli* were transfected with the plasmid for either EcoRI^{E111Q}-3xFLAG or LacI-3xFLAG and grown overnight on LB (Sigma-Aldrich) agar plates. A suspension culture from a single colony was grown to 1.3 OD_{600} while shaking at 37°C and used to inoculate 200 mL culture which was then grown to 0.5 OD_{600} before induction. Protein expression was induced with 2.5 mM IPTG (Sigma-Aldrich) and allowed to continue for 4 hours. Cells were pelleted by centrifugation at 3000xg for 10 minutes then resuspended in column buffer (20 mM Tris-HCl pH 8.5, 500 mM NaCl) supplemented with a protease inhibitor tablet (Pierce, Thermo Scientific), 1 mg/mL lysozyme (Calbiotech, Inc.) and stored on ice for 20 minutes. Cells were lysed with 3 cycles of alternating freeze-thaw using dry ice and a 42°C water bath followed by 5 minutes of alternating sonication, using a probe sonicator, and cooling on ice to prevent heat damage to the protein. Protein was pelleted by centrifugation at 18000xg for 30 minutes at 4°C. The supernatant was purified by column affinity chromatography using chitin resin (New England Biolabs, Inc) and, after rinsing thoroughly with column buffer, eluted with thiol

buffer (column buffer + 50 mM DTT). The buffer was exchanged for storage buffer (40 mM Tris pH 7.5, 300 mM NaCl, 0.1 mM EDTA, 10 mM 2-mercaptoethanol) via dialysis and spin concentrated using a Spin-X UF concentrator (Corning®).

Electrophoretic mobility shift assay

Complimentary 40 bp DNA oligomers, one of which was biotinylated, containing a single protein binding site were annealed to create the double stranded substrate to measure dissociation constant of EcoRI^{E111Q}-3xFLAG and LacI-3xFLAG.

For EcoRI^{E111Q}-3xFLAG, the DNA fragment contained the binding sequence GAATTC. The binding experiment was carried out in EcoRI binding buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM $MgCl_2$, 0.2 mg/mL bovine serum albumin (BSA)) with 50 ng/µL Poly (dI-dC).

For LacI-3xFLAG, the binding sequence used was AATTGTTATCCGCTCACAATT, again centrally located on the DNA molecule. The binding reaction was carried out in LacI binding buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.2 mg/mL BSA) with 50 ng/µL Poly (dI-dC).

The protein-DNA complex was run in a 6% PAGE gel and labeled using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific) and the provided protocol. The chemiluminescence was visualized using a gel imager provided by the Bogdanove Laboratory at Cornell University.

Protein binding experiments

80 nM EcoRI^{E111Q}-3xFLAG was incubated with 100 pM lambda DNA in 50 µL of EcoRI binding buffer for 20 minutes prior to dilution in imaging buffer to 1mL

final volume and injection into chamber. After allowing the protein-DNA complexes to anchor themselves to the bilayer and the double tether anchor for 30 minutes, 700 µL of 1.6 µg/mL anti-FLAG was incubated for 20 minutes then flushed out with imaging buffer. Finally, 700 µL of 400ng/mL Alexa-647 secondary antibody was incubated for 20 minutes and flushed with imaging buffer prior to imaging.

10 nM LacI-3xFLAG was incubated with 350 pM DNA 2xLacO-lambda DNA in 25 µL of LacI binding buffer for 20 minutes prior to dilution in imaging buffer to 1 mL final volume. The remaining steps are the same as was used for EcoRI^{E111Q}-3xFLAG.

TAL1535-SNAP was incubated with SNAP-Alexa 647 and excess dye was removed. The labeling rate was approximately 5%.30-1500 nM TAL1535-Alexa 647 (~1-100 nM labeled TAL protein) was incubated with 10-200 pM TAL1535-lambda DNA in 100 µL of TAL binding buffer for 25 minutes prior to dilution in 1 mL final volume and injection into chamber.

REFERENCES

1. Cuculis, L., et al., *Direct observation of TALE protein dynamics reveals a two-state search mechanism.* Nat Commun, 2015. **6**: p. 7277.
2. Cermak, T., et al., *Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting.* Nucleic Acids Res, 2011. **39**(12): p. e82.
3. Shashkova, S. and M.C. Leake, *Single-molecule fluorescence microscopy review: shedding new light on old problems.* Biosci Rep, 2017. **37**(4).
4. Cuculis, L. and C.M. Schroeder, *A Single-Molecule View of Genome Editing Proteins: Biophysical Mechanisms for TALEs and CRISPR/Cas9.* Annu Rev Chem Biomol Eng, 2017. **8**: p. 577-597.
5. Visnapuu, M.L., et al., *Parallel arrays of geometric nanowells for assembling curtains of DNA with controlled lateral dispersion.* Langmuir, 2008. **24**(19): p. 11293-9.
6. Bogdanove, A.J. and D.F. Voytas, *TAL effectors: customizable proteins for DNA targeting.* Science, 2011. **333**(6051): p. 1843-6.
7. Graneli, A., et al., *Organized arrays of individual DNA molecules tethered to supported lipid bilayers.* Langmuir, 2006. **22**(1): p. 292-9.
8. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nat Methods, 2008. **5**(6): p. 507-16.
9. Soltani, M., et al., *Nanophotonic trapping for precise manipulation of biomolecular arrays.* Nat Nanotechnol, 2014. **9**(6): p. 448-52.

10. Plenat, T., et al., *High-throughput single-molecule analysis of DNA-protein interactions by tethered particle motion.* Nucleic Acids Res, 2012. **40**(12): p. e89.

11. Palma, M., et al., *Selective biomolecular nanoarrays for parallel single-molecule investigations.* J Am Chem Soc, 2011. **133**(20): p. 7656-9.

12. Kaykov, A., et al., *Molecular Combing of Single DNA Molecules on the 10 Megabase Scale.* Sci Rep, 2016. **6**: p. 19636.

13. Greene, E.C., et al., *DNA curtains for high-throughput single-molecule optical imaging.* Methods Enzymol, 2010. **472**: p. 293-315.

14. Lee, J.Y., et al., *Single-molecule imaging of FtsK translocation reveals mechanistic features of protein-protein collisions on DNA.* Mol Cell, 2014. **54**(5): p. 832-43.

15. Robison, A.D. and I.J. Finkelstein, *Rapid prototyping of multichannel microfluidic devices for single-molecule DNA curtain imaging.* Anal Chem, 2014. **86**(9): p. 4157-63.

16. Fazio, T., et al., *DNA curtains and nanoscale curtain rods: high-throughput tools for single molecule imaging.* Langmuir, 2008. **24**(18): p. 10524-31.

17. Fazio, T.A., et al., *Fabrication of Nanoscale "Curtain Rods" for DNA Curtains Using Nanoimprint Lithography.* J Vac Sci Technol A, 2009. **27**(6): p. 3095-3098.

18. Gorman, J., et al., *Nanofabricated racks of aligned and anchored DNA substrates for single-molecule imaging.* Langmuir, 2010. **26**(2): p. 1372-9.

19. Terakawa, T., et al., *Sequential eviction of crowded nucleoprotein complexes by the exonuclease RecBCD molecular motor.* Proc Natl Acad Sci U S A, 2017. **114**(31): p. E6322-E6331.

20. Finkelstein, I.J., M.L. Visnapuu, and E.C. Greene, *Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase.* Nature, 2010. **468**(7326): p. 983-7.

21. Moscou, M.J. and A.J. Bogdanove, *A simple cipher governs DNA recognition by TAL effectors.* Science, 2009. **326**(5959): p. 1501.

22. Bogdanove, A.J., S. Schornack, and T. Lahaye, *TAL effectors: finding plant genes for disease and defense.* Curr Opin Plant Biol, 2010. **13**(4): p. 394-401.

23. Boch, J., et al., *Breaking the code of DNA binding specificity of TAL-type III effectors.* Science, 2009. **326**(5959): p. 1509-12.

24. Boch, J., *TALEs of genome targeting.* Nat Biotechnol, 2011. **29**(2): p. 135-6.

25. Miyanari, Y., C. Ziegler-Birling, and M.E. Torres-Padilla, *Live visualization of chromatin dynamics with fluorescent TALEs.* Nat Struct Mol Biol, 2013. **20**(11): p. 1321-4.

26. Mahfouz, M.M., et al., *De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks.* Proc Natl Acad Sci U S A, 2011. **108**(6): p. 2623-8.

27. Gao, X., et al., *Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers.* Nucleic Acids Res, 2014. **42**(20): p. e155.

28. Zhang, S., H. Chen, and J. Wang, *Generate TALE/TALEN as Easily and Rapidly as Generating CRISPR.* Mol Ther Methods Clin Dev, 2019. **13**: p. 310-320.

29. Mussolino, C., et al., *A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity.* Nucleic Acids Res, 2011. **39**(21): p. 9283-93.

30. Cuculis, L., et al., *TALE proteins search DNA using a rotationally decoupled mechanism.* Nat Chem Biol, 2016. **12**(10): p. 831-7.

31. Cho, H., et al., *Low mobility of phosphatidylinositol 4,5-bisphosphate underlies receptor specificity of Gq-mediated ion channel regulation in atrial myocytes.* Proc Natl Acad Sci U S A, 2005. **102**(42): p. 15241-6.

32. Ellenberg, J., et al., *Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis.* J Cell Biol, 1997. **138**(6): p. 1193-206.

33. Petrosyan, R., *Improved approximations for some polymer extension models.* Rheologica Acta, 2017. **56**(1): p. 21-26.

34. Wang, M.D., et al., *Stretching DNA with optical tweezers.* Biophys J, 1997. **72**(3): p. 1335-46.

35. Bustamante, C., Z. Bryant, and S.B. Smith, *Ten years of tension: single-molecule DNA mechanics.* Nature, 2003. **421**(6921): p. 423-7.

36. Forde, G.M., et al., *LacO-LacI interaction in affinity adsorption of plasmid DNA.* Biotechnol Bioeng, 2006. **95**(1): p. 67-75.

CHAPTER 3

STOICHIOMETRIC ANALYSIS OF PROTEIN COMPLEXES BY CELL FUSION AND SINGLE MOLECULE IMAGING

The composition, stoichiometry, and interactions of supramolecular protein complexes are a critical part of biological function. Several techniques have been developed to study molecular interactions and quantify subunit stoichiometry at the single molecule level; however, these typically require artificially low expression levels to achieve single molecule imaging or use of detergent isolation of complexes that may perturb native subunit interactions. Here we present an alternative approach where protein complexes are assembled at physiological concentrations and subsequently diluted in situ for single-molecule level observations while preserving them in a near-native cellular environment. We show that coupling this in situ dilution strategy with single molecule techniques, such as in vivo Fluorescence Correlation Spectroscopy (FCS), bleach step counting for quantifying protein complex stoichiometry, and two-color single molecule colocalization, improves the quality of data obtained using these single molecule fluorescence methods. We call the technique Single Protein Recovery After Dilution (SPReAD) and demonstrate that it is a simple and versatile means of extending the concentration range of single molecule measurements into the cellular regime while minimizing potential artifacts and perturbations of the functional protein complex stoichiometry.

INTRODUCTION

Dynamic networks of protein interactions underlie much of cell biology. A key goal of biomedical science is to understand the nature of these interactions and elucidate how they change in response to various extracellular cues. Screening methods, such as yeast-two hybrid analysis or phage display, are useful for identifying potential binding partners in a high-throughput manner, but generally ignore biological context [1]. Ensemble approaches that rely on co-immunoprecipitation or fluorescence spectroscopy can more accurately capture interactions within the cellular environment and can be used to examine changes that occur in response to external stimuli [1, 2]. However, these bulk ensemble averaged measurements yield little information about the stoichiometry of subunits within complexes. In contrast, single molecule methods have the sensitivity to probe single protein complexes and quantitatively report on their individual architecture.

Early uses of single-molecule fluorescence for subunit counting relied on artificially low expression levels in order to resolve individual protein complexes [3]. While this method was effective, non-physiological protein concentrations during assembly may shift binding equilibria and alter normal stoichiometry. More recently, a single-molecule pull-down (SiM-Pull) approach has been developed that allows complexes to be assembled at native expression levels, extracted into a cell lysate, and then captured on an antibody-coated slide for single-molecule imaging [4]. Antibody concentrations and lysate dilutions can be tuned to maintain single-molecule resolution without compromising intracellular assembly

conditions. Although SiM-Pull has been used to successfully measure the subunit stoichiometry of membrane receptors, mitochondrial proteins, virus replication initiation, nuclear export complexes, and signaling complexes, the use of detergents for isolation and subsequent wash steps may affect the integrity of assemblies, and therefore, the physiological relevance of the data [5].

Here, we introduce a simple detergent-free method to examine single protein complexes assembled at normal physiological concentrations in a near-native environment. Two cell populations—one containing a protein complex of interest and the other lacking it—are co-plated on a coverslip and fused into large syncytia. Protein diffusion within these syncytia results in a uniform dilution of labelled complexes permitting their examination at reduced concentrations. Dilution factors are controllable by varying the plating ratio between expressing and non-expressing cells. This dilution factor can be made large enough that single membrane protein complexes may be resolved for use in stepwise photobleaching and brightness analysis, 2D membrane fluorescence fluctuation spectroscopy (FCS), two-color single molecule colocalization or single molecule FRET experiments. Cytosolic proteins are also subject to this dilution, enabling high-quality in vivo FCS and FCCS data to be collected. We call our approach Single Protein Recovery After Dilution (SPReAD) because it yields concentrations suitable for single molecule imaging after physiological oligomer assembly.

RESULTS

Formation of large syncytia using an inducible VSVG

A stable cell line with conditional expression of vesicular stomatitis virus G

protein (VSVG) was created and used to initiate controlled cell fusion between the VSVG-expressing cell line and cells expressing a labeled protein of interest. VSVG is a well-characterized fusogen that can be reversibly activated by a short pH drop [6]. To dilute protein complexes for stoichiometric analysis, doxycycline-inducible VSVG-expressing USOS cells are mixed with cells expressing the target protein, typically at a 10:1 ratio (VSGV cells to target cells), and incubated at pH 5.5 for 1 to 5 minutes. After activating VSVG in a confluent monolayer of the mixed cell culture, we observed rapid formation (<1 hour) of massive syncytia and diffusion of labeled proteins producing a uniform distribution in which single molecules (adrenergic receptors in this case) can clearly be identified for analysis (Figure 3.1a and b, right). The signal-to-noise level in the resulting images was equivalent to that which we measured using single molecule pull-downs of detergent isolated protein from the same cells (Figure 3.1b, left). These results suggest that substantial dilution factors may be attained in time intervals comparable to the handling time for cell lysate preparation, implying that the two approaches have similar bounds for cases where oligomerization is non-covalent and transient. However, SPReAD has the advantage of not requiring a detergent isolation step and the initial intracellular complex formation is carried out under normal cellular conditions before fusion.

The dynamics of cell fusion and long-term viability of syncytia was visualized using continuous bright-field imaging. We observed a loss of cell

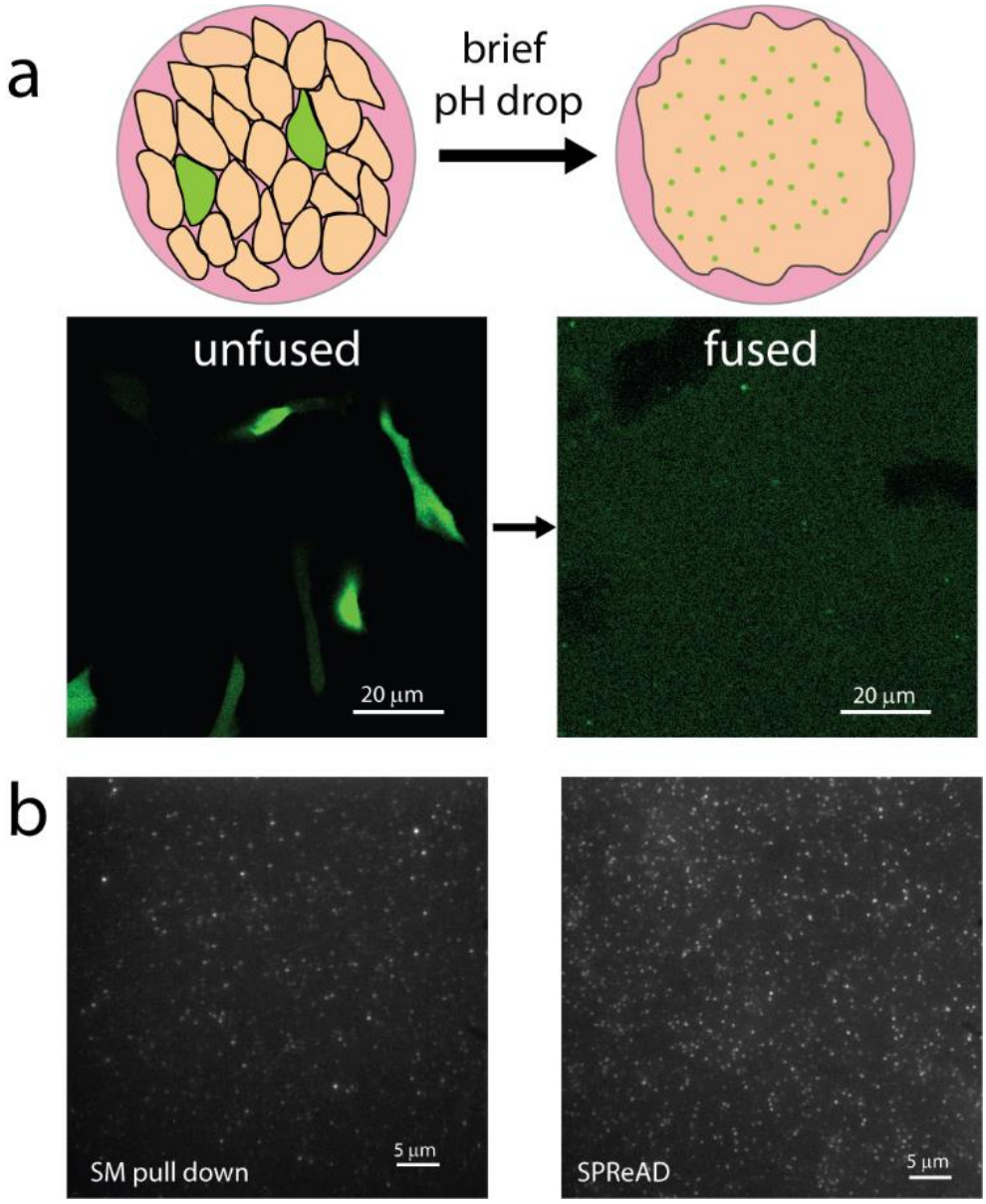


Figure 3.1 Use of Single Protein Recovery After Dilution (SPReAD) for Single Molecule Imaging

(a) Cells expressing a labeled protein-of-interest (green) are co-plated with a stable U2OS cell line (orange) which express VSVG after doxycycline activation. A brief incubation in low pH (5.5-6.0) buffer initiates membrane fusion, after which protein complexes diffuse out of their parent cells into the larger syncytium. (b) mNG-β_2AR protein complexes prepared for single molecule imaging by detergent isolation and biotin-streptavidin pull-down (left) and mNG-β_2AR protein complexes in the syncytium membrane after VSVG-mediated fusion (right).

boundaries approximately 40 minutes after pH drop. For the next ~4-5 hours, the syncytium remained bound to the coverslip and displayed few morphological changes. Thereafter, adhesion was slowly lost over the course of 12 hours, and at 20 hours, concerted cell death was observed. This suggests a 4-5 hour period during which cells are fused, but otherwise minimally perturbed. Cells may be imaged live during this window or fixed for later observation. Local spreading by diffusion, of both cytosolic and membrane proteins occurs rapidly, and single molecule compatible levels are reached in 20-30 minutes. Therefore we conclude that SPReAD allows oligomerization to be studied while preserving the biophysical environments of membrane and cytosolic protein complexes.

Large-scale fusion was possible in all mammalian cell lines tested as expected due to the broad tropism of VSVG. This provides an additional means of experimental control by which cofactors can be included or excluded by choice of cell type.

The formation of syncytia is a key step in the development of various mammalian tissues including bone, muscle, and placenta [7]. In these cases, cell fusion is well regulated and part of the normal developmental program. Cell fusion can also play a role in the progression of disease. Many enveloped viruses trigger fusion between an infected cell and its neighbors resulting in new and abnormal hybrids. Accidental cell fusion, both due to viral infection and otherwise, has also been implicated in cancer, where polyploid cells display high levels of chromosomal instability and may acquire tumorigenic phenotype [8]. These natural examples of cell fusion suggest that large syncytia, at least during the first few

hours after fusion, constitute a minimally perturbative system in which proteins preserve physiological arrangement and behavior.

Dilution of labelled cytosolic proteins by cell fusion improves *in vivo* Fluorescence Correlation Spectroscopy

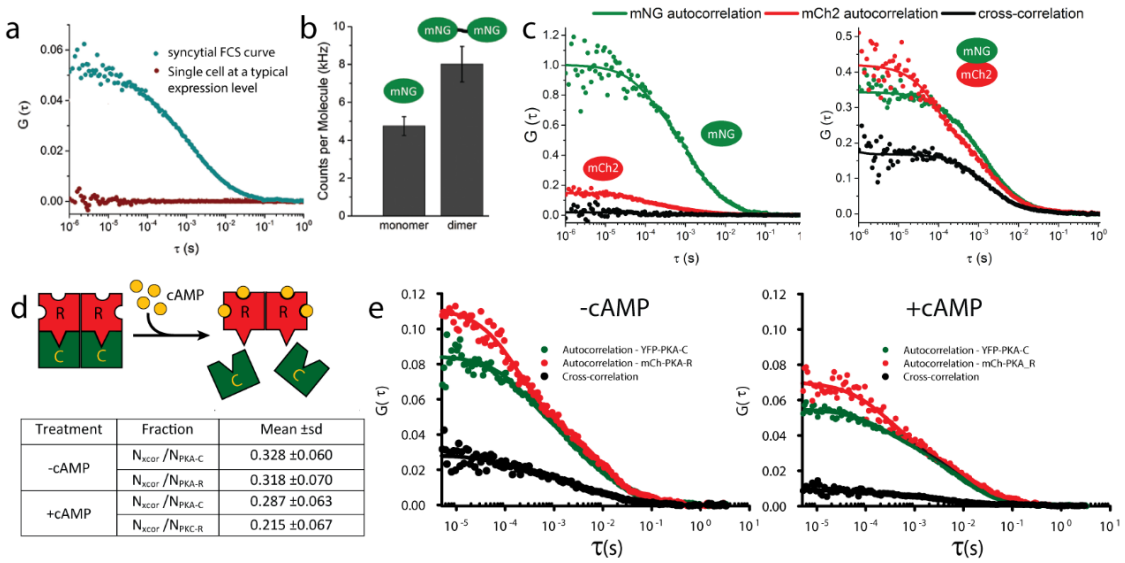
Since its invention in the 1970s, fluorescence correlation spectroscopy (FCS) has become a valuable tool for investigation of molecular transport and interactions [9]. Autocorrelation analysis provides information about diffusion, per-particle brightness, and local concentrations, while two-color fluorescence cross-correlation spectroscopy (FCCS) can probe molecular associations [10]. FCS and FCCS can be used inside living cells, but cellular proteins are typically expressed at intracellular concentrations outside of the working range for FCS studies. Furthermore, standard FCS and FCCS fitting models assume an infinite pool of diffusive species such that molecular motions are unconstrained and photobleaching is inconsequential. However, this is hardly the situation within the cellular environment and is known to lead to artifacts [11]. Cell fusion is a promising means to address both of these limitations as concentrations can be arbitrarily adjusted to fall inside the optimal FCS range, and the relatively large size of the syncytium serves to alleviate the effects of constrained motion and photobleaching-induced concentration reduction that can occur during live-cell FCS. For cell fusion to function as a dilution strategy, protein complexes must be sufficiently mobile to diffuse out of their parent cells into the larger syncytium. Proteins confined to specific organelles or stably tethered to cytoskeletal components may fail to satisfy this criterion; however, many transcription factors

and signaling complexes have mobile cytoplasmic fractions and are candidates for cell fusion-based dilution and single molecule analysis. The kinetics of syncytium formation and protein mobility determine the optimal timeframes for imaging and fixation after fusion is initiated. Time-lapse imaging revealed that membrane fusion was immediate and synchronized across the imaging vessel, with cytosolic proteins beginning to escape their parent cells within 2 minutes of pH. The initially heterogeneous fluorescence distribution was continually reshaped by diffusion until reaching a uniform steady-state level after ~30 minutes. The equilibration time will depend on the size of protein complexes being studied, their interactions with static cellular components, experimental conditions, and the ratio of expressing and non-expressing cells. Overall, the kinetics of cell fusion and protein redistribution provide two possible modes of measurement. Measurements made in the non-equilibrium stage will most accurately report on the stoichiometry of weakly interacting protein complexes because assemblies have less time to dissociate before recording. However, concentration measurements during this stage may be less reliable due to the heterogeneity across the imaging dish. In contrast, equilibrium stage measurements can be used to estimate the original intracellular concentrations (based on a known co-plating ratio), but may provide less accurate stoichiometric measurement of complexes with the fastest dissociation rates. Overall, this flexibility renders SPReAD as a versatile method for quantification of both oligomeric state and cellular expression levels.

To determine the range of dilutions possible, non-fluorescent VSVG-expressing cells were mixed with cells stably expressing mNeonGreen (mNG) at

various plating ratios. After fusion, the fluorescence signal per unit area dropped in proportion to the plating ratio. Absolute numbers for syncytial concentrations were obtained by fluorescence correlation spectroscopy (FCS) and showed a similar trend, deviating only higher concentrations where FCS-based quantification is unreliable. We found that fusion-based dilution could be used to adjust cytoplasmic levels of an expressed protein over of range ~two orders of magnitude. Importantly, larger dilutions brought cytosolic levels down to the sub-100 nM range where correlation spectroscopy functions best.

To explore the benefits of using SPReAD for intracellular FCS measurements, we compared FCS curves obtained in unfused cells with those in syncytia (Figure 3.2a). In cells, transient mNG expression from a CMV promoter often failed to produce usable autocorrelation curves due to the high cytosolic concentration of labeled protein following transfection. FCS data is typically not useable when fluorophore levels exceed ~1 μM, which is well within the range of normal intracellular protein concentrations. In practice, one either picks cells with low enough expression to obtain useable correlation curves or carries out whole cell photobleaching to reduce the fluorescent species concentration to FCS-compatible levels [12]. Both of these options have clear biological drawbacks – either biasing the results by selecting only cells with low levels of expression or phototoxicity from the bleaching method. We found that autocorrelations in unfused cells had an average dwell time of 2.2 ± 1.3 ms corresponding to a diffusion coefficient of 10 ± 5.8 $\mu m^2/s$. We attribute the large deviations in the measured values (~50%) from overall poorer data quality due to the



Treatment	Fraction	Mean ±sd
-cAMP	$N_{xcor}/N_{PKA\text{-}C}$	0.328 ±0.060
	$N_{xcor}/N_{PKA\text{-}R}$	0.318 ±0.070
+cAMP	$N_{xcor}/N_{PKA\text{-}C}$	0.287 ±0.063
	$N_{xcor}/N_{PKC\text{-}R}$	0.215 ±0.067

Figure 3.2 SPReAD Improves *in vivo* Fluorescence Correlation Spectroscopy (FCS) Measurements

(a) FCS curves from syncytia are of uniform high quality since the concentrations can be set to FCS-compatible levels, compared to expression levels generally found in single transformed cells. (b) Brightness based on counts per molecule can be used to discriminate between monomeric and dimeric species in the cytoplasm of large syncytia and is useful for studying the stoichiometry of a single component within oligomers. (c) Cross-correlation spectroscopy is useful for studying heteromeric interactions. mNeonGreen and mCherry produce an appreciable cross-correlation (black line) when covalently joined (right) but not when co-transfected (left). In both cases, auto-correlations for each protein are clearly discernable. (d and e) FCCS in syncytia can be used to study functional differences in protein oligomerization. Protein Kinase A regulatory and catalytic subunits form complexes in the baseline state, repressing activity. Stimulation of adenylyl cyclase generates cAMP, causing PKA dissociation and increased activity. Table values in (d) represent data from 13 experiments.

measurements being made at higher than ideal fluorophore concentrations and the altered mobility near bounding membranes or organelles with in the single cells. We often saw artifacts due to photobleaching which manifest as a change in G(0) over time. In comparison, dwell times and G(0) values from syncytial data FCS curves showed much less variation due to the larger homogenous pool of diffusing fluorophores. Syncytial FCS curves yielded dwell times and diffusion coefficients (1.2 ± 0.1 ms; 13 ± 1.1 $\mu m^2/s$) similar to the unfused cells, but with much less variation. Although diffusion coefficients are a relatively weak indicator of changes in oligomerization state or ligand binding ($D \approx M^{1/3}$), the substantial reduction in variation measurement variation afforded by SPReAD can improve the FCS methodology for these types of investigations.

Brightness analysis and two-color fluorescence cross-correlation spectroscopy are two valuable methods for studies of protein-protein interactions within the cellular environment [10, 13]. To evaluate these techniques in conjunction with cell fusion, we compared measurements made with covalent dimers of fluorescent proteins to the corresponding monomeric proteins. mNG dimers were found to be 1.7 times brighter than monomers (Figure 3.2b). Assuming quenching effects are minimal, this suggests a maturation efficiency of 80-85% for mNG, which is on par with that of other green/yellow fluorescent proteins. Two color fluorescence cross-correlation spectroscopy (FCCS) of mNeonGreen-mCherry2 covalent dimers yielded a 58% dimer population, while a co-transfection of the monomeric proteins showed negligible cross-correlation amplitude (Figure 3.2c). In addition to brightness and cross-correlation analyses,

other methods such as photon counting histograms, dwell time distributions, photon anti-bunching, and single-molecule FRET have been used to examine oligomerization states and could be aided by SPReAD sample preparation.

Next, we used syncytial two-color FCCS to study the oligomerization of protein kinase A (PKA), a Ser/Thr kinase that functions in the cAMP-dependent pathway of GPCR signaling. Upon GPCR activation, adenylyl cyclase catalyzes the conversion of ATP into cAMP causing PKA regulatory subunits to dissociate from catalytic subunits which are then free to phosphorylate downstream targets. Syncytial FCS of YFP-tagged catalytic subunits and mCherry-tagged regulatory subunits revealed a significant cross-correlation indicating functional repression in the baseline state (Figure 3.2d, e). Upon stimulation with the adenylyl cyclase activator forskolin and the phosphodiesterase inhibitor IBMX, cross-correlation amplitudes decreased, reflecting cAMP-induced dissociation of subunits and mirroring previous efforts using FCS in live cells or SiM-Pull with cell lysates [4, 14]. SPReAD increases the usefulness and robustness of FCS and FCCS for cell based measurements by allowing for target complex formation at more physiological intracellular concentrations and by mitigating complicating effects from confined cellular volumes.

Single molecule imaging of membrane protein complexes

Most membrane proteins are freely mobile in two dimensions unless tethered to intracellular actin. Membrane residing protein complexes are of significant interest to biomedical research, representing 23% of all ORFs in the human genome and being the target of >60% of pharmaceutical drugs [15]. The

biomedical significance of membrane receptors has motivated comprehensive investigation of their basic structures and mechanisms of action. Oligomerization is known to play a role in the function of all major receptor types (metabotropic, ionotropic, and tyrosine kinases), and thus, considerable effort has been made to elucidate their interaction profiles. From a single-molecule perspective, subunit counting in oocytes has been the widely used approach, with many receptors being studied after controlled mRNA injection to limit receptor levels [16]. However, the concentration-dependence of oligomerization may be at odds with the sub-physiological expression levels employed in this technique. Furthermore, cell type specific post-translational modifications, occurring in the Golgi and ER required for native oligomer formation, may be lacking. We demonstrate that cell fusion, when combined with single molecule imaging, lifts this restriction and allows single molecule imaging after physiological assembly of receptor complexes in a cell type required by any specific biological constraints.

We undertook a series of experiments designed to demonstrate the unique utility of SPReAD for single molecule imaged based measurements of membrane protein stoichiometry. In these experiments, we examined differences between in the results obtained in SPReAD preparations and single molecule pull-downs to determine if detergent isolations had a notable effect, as well as investigations in the oligomeric state of several well-studied membrane protein complexes.

Beta-2 adrenergic receptor stoichiometry

The β2 adrenergic receptor (ADRβ2) is a prototypical G protein coupled receptor known to form both homodimers as well as heterodimers with other

members of the GPCR family [17]. We used a mNG-tagged derivative of ADRβ2 expressed in U2OS cells co-plated with VSVG-expressing neighbors to dilute membrane receptors from the initial high expression levels. After cell fusion and incubation at 37°C for 1 hour, individual receptor complexes were clearly discernible and mobile within the plasma membrane, displaying similar kinetics to measurements made in living cells (Figure 3.1b). Single particle tracking of mNG-ADRβ2 also confirmed this observation. The receptor concentration distribution was more heterogeneous across the syncytium than the distribution which we observed with cytosolic proteins even ~1 hour after fusion due to the slower diffusion rate for proteins in the membrane compared to the cytoplasm. However, there were still many fields of view with uniform sparse distributions ideal for single molecule imaging and analysis (Figure 3.1b, right panel).

Syncytia were fixed with paraformaldehyde to immobilize receptor complexes and facilitate stoichiometry determination by stepwise photobleaching and two-color single molecule co-localization methods. mNG-ADRβ2 puncta showed distinct bleach steps (Figure 3.3a). Analysis of the receptor population revealed that ADRβ2 was evenly distributed between monomeric and dimeric states, with 25% of photobleaching traces showing two bleach steps (Figure 3.3b, fourth bar group), signifying a 36% dimer population after accounting for mNG's maturation efficiency. In order to measure the effects of DDM, a detergent commonly used in cell lysis for pull-down experiments, we carried out single-molecule pull-down (SiM-Pull protocol) experiments on the same mNG-ADRβ2 expressing cells and found that dimer fractions differed significantly from what was

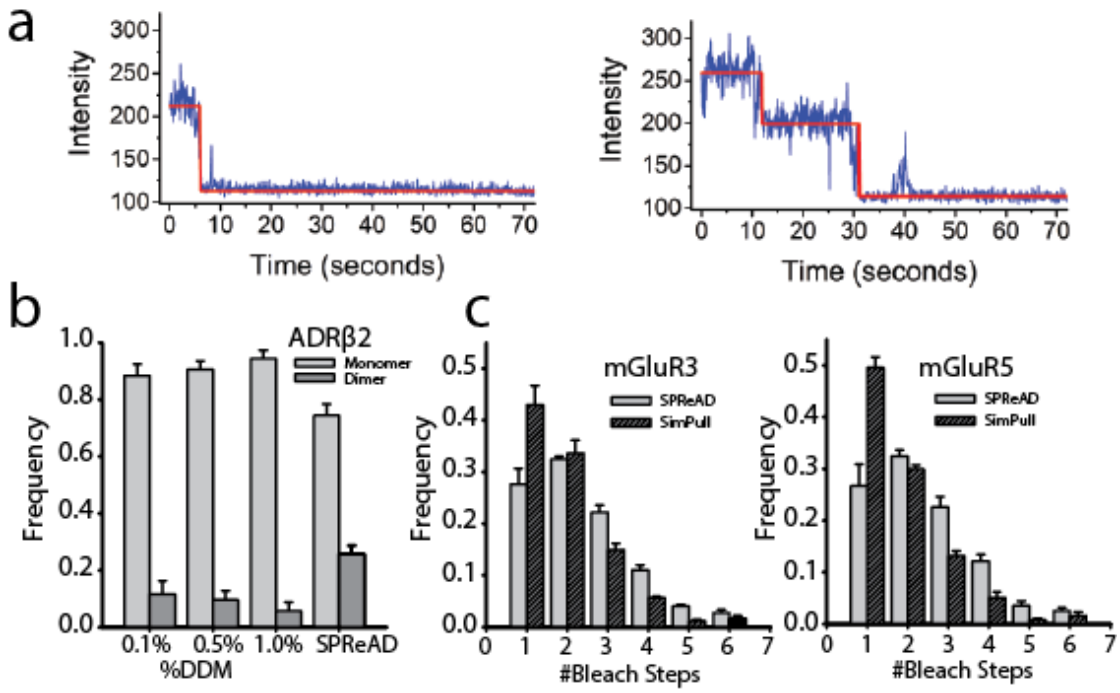


Figure 3.3 SPReAD for Single Molecule Imaging Avoids Potential Artifacts of Detergent Isolation

(a) Typical monomer (left) and dimer (right) mNG-ARβ2 bleach step traces obtained from the SPReAD-prepared samples. (b) The elimination of detergent isolation artifacts was demonstrated by measuring the mNG-ARβ2 dimer to monomer ratio from pull-down experiments using mNG-ARβ2 isolated at three different detergent concentrations (first three groups of bars) and from SPReAD preparations (fourth bar group). Data is the mean ± SEM, n = 3. These experiments showed a significantly larger fraction of mNG-ARβ2 dimers using SPReAD, which we take to be a more accurate measure of the physiological dimer ratio. (c) Comparison of metabotropic glutamate receptor (mGluR3 and mGluR5 – both known to form homodimers) complex stoichiometry in samples prepared via SPReAD, or via lysis for SiM-Pull (mean ± SEM, n = 3). We demonstrate that detergent-isolated single molecule pull-down experiments show a larger monomer fraction relative to SPReAD, which we interpret as an effect of the isolation treatment or subsequent wash steps on complex stability.

observed with SPReAD (Figure 3.3b, first three bar groups). When performing a standard detergent isolation and single molecule pull-down procedure the dimer to monomer ratio was never higher than 10%.

Metabotropic glutamate receptor stoichiometry

The metabotropic glutamate receptors mGluR3 and mGluR5 are known to function as covalently bound homodimers via a cysteine bridge assembled in the ER prior to membrane trafficking [18]. We generated stable HEK293T cell lines constitutively expressing a mNeonGreen labeled metabotropic glutamate receptor, either mNG-mGluR3 or mNG-mGluR5, and co-plated them with VSVG-expressing U2OS cells. Fusion was induced and syncytia were formed and fixed with paraformaldehyde as described above. After fusion, individual receptor complexes were able to be resolved in many areas of the dish, demonstrating that fusion can occur between different cell types and that proteins are capable of diffusing throughout the heterogeneous membrane of these syncytia.

Stepwise photobleaching quantification was performed on the syncytia and we found that 33.5 ± 0.5% and 33.8 ± 0.8% of the traces showed two bleach steps for mNG-mGluR3 and mGluR5, respectively, confirming that complexes formed prior to membrane trafficking were preserved during the fusion process (Figure 3.3c). Single-molecule pull-down experiments were then performed on the same mNG-mGluR3 and 5 expressing HEK293T cell lines using 1% DDM in the lysis buffer. These yielded slightly lower fractions of dimers for both mGluRs, with 29.9 ± 0.8% of mNG-mGluR3 and 28.1 ± 0.4% of mNG-mGluR5 traces demonstrating two-step bleaching (Figure 3.3c).

Detection of higher-order oligomeric membrane protein complexes – CRAC channel subunit Orai1 stoichiometry

We also examined the subunit stoichiometry of Orai1, a calcium-selective ion channel that forms the central pore of the calcium release-activated channel. The functional stoichiometry is currently unresolved, with claims of either a tetrameric or a hexameric configuration [19-21]. Using SPReAD, we found that most Orai1 puncta bleached in 1-6 steps (Figure 3.4), which tends towards the hexametric structure model proposed in [20], especially when correction of the raw bleach step distribution for an estimated dark fraction of mNeonGreen is taken into account. Assuming a 20% misfolded protein, the corrected distribution has a weighted mean of 5 Orai1 subunits per complex. Although not fully hexameric on average, this result may be due to the interference by the mNG moiety and/or the presence of unlabeled Orai1 in mammalian cells. Targeted knockdown of endogenous proteins or careful choice of cell lines may be used to refine understanding of physiologically-relevant oligomerization in specific tissue types.

Ligand-dependent oligomerization of epidermal growth factor receptor

One of the primary evolutionary advantages conferred by oligomerization is the development of new modes of regulatory control of protein activity. Allosteric oligomerization is known to play a role in the mechanisms of both metabotropic receptors and receptor tyrosine kinases, with extracellular ligands modulating the formation of dimers or higher order structures. Monomer-oligomer transitions can prime receptors for downstream signaling events, such as posttranslational modifications or the recruitment of adaptor proteins. In order to evaluate the use

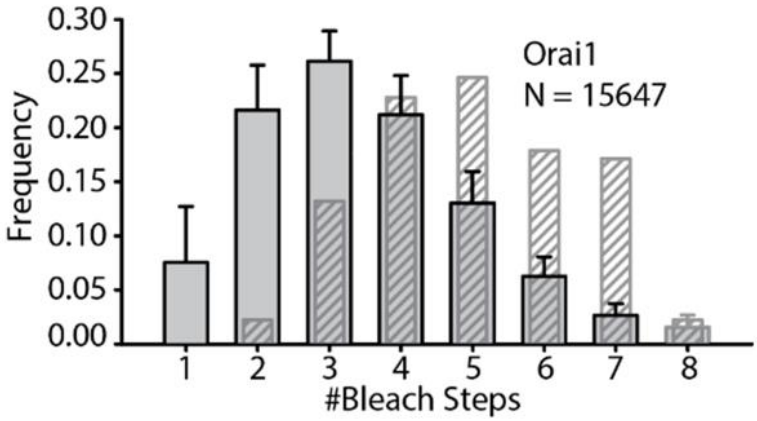


Figure 3.4 Application of SPReAD to Detect Higher-Order Oligomer Membrane Protein Complexes

The functional form of the CRAC channel continues to be a contested issue, with multiple sources claiming the Orai1 subunits adopt either a tetrameric or a hexameric configuration. The raw bleach step data from 15647 fluorescent spots analyzed from SPReAD syncytia made from Orai1 expressing cells yielded a weighted mean step number of 3.5 (gray bars). Hatched bars represent an estimate of the actual subunit fractions assuming a 0.8 fluorescent fraction for mNG and indicate a weighted mean subunit number of 5.0.

of SPReAD to detect ligand-dependent multimerization, we examined epidermal growth factor receptors—a member of the receptor tyrosine kinase family whose abnormal signaling has been implicated in a number of human cancers [22]. The canonical model for receptor activation asserts that EGFR is monomeric in the plasma membrane prior to stimulation, whereupon it is driven to dimerize by the binding of its cognate ligands, resulting in autophosphorylation of tyrosine residues on its cytoplasmic tail and recruitment of specific effector proteins (Figure 3.5a).

Although the EGFR pathway has been extensively studied using both bulk and single molecule approaches, there are still open questions about receptor oligomerization. There is increasing evidence that pre-formed dimers of EGFR exist on the cell surface prior to ligand stimulation and that EGFR is capable of forming higher-order oligomers that may function in receptor activation [23-25]. To examine each of these possibilities, we expressed an mNeonGreen-tagged EGFR (mNG-EGFR) on the cell surface and performed stepwise photobleaching measurements in fixed syncytia. Even in its baseline state, EGFR was found to be significantly dimeric, with 29% of traces bleaching in two steps (Figure 3.5b). Upon stimulation with EGF, this dimer fraction increased substantially and some higher-order oligomers (trimers and tetramers) were observed. Together, these results support a model where at least some EGFR signaling is accomplished by conformational changes in pre-formed dimers and trans-activation by higher-order oligomers. The use of SPReAD to study ligand-dependent oligomerization of EGFR validates its potential for studying transient interactions. Future studies may aim to visualize the recruitment of specific factors such as GRB2 or PLCγ1, that

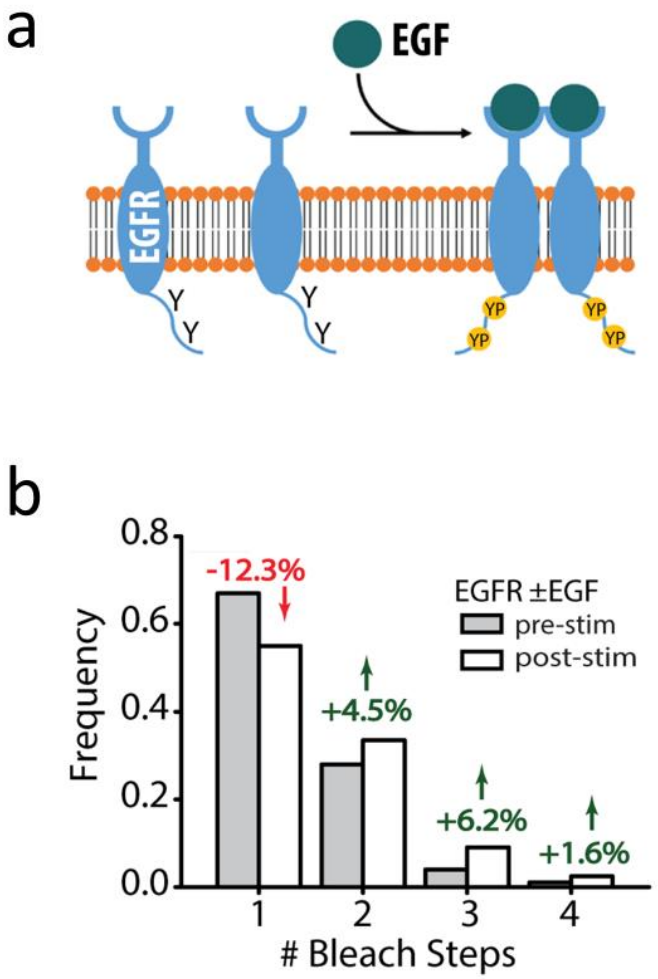


Figure 3.5 Utilizing SPReAD to Measure Stoichiometric Changes Due to Ligand Binding

(a) Epidermal growth factor binding stimulates EGFR dimerization and autophosphorylation of tyrosine residues on EGFR's cytoplasmic tail, leading to recruitment of downstream signaling proteins. (b) SPReAD bleach step histograms of EGFR oligomerization before (gray bars) and after (white bars) EGF stimulation. Although EGFR is largely monomeric prior to growth factor addition, there is a substantial dimer fraction as well. After stimulation, the dimer and higher-order oligomers fractions increase, while the monomer fraction drops.

relay extracellular signals to downstream effectors, or the impact of pharmacological agents or EGFR mutations on oligomerization propensity.

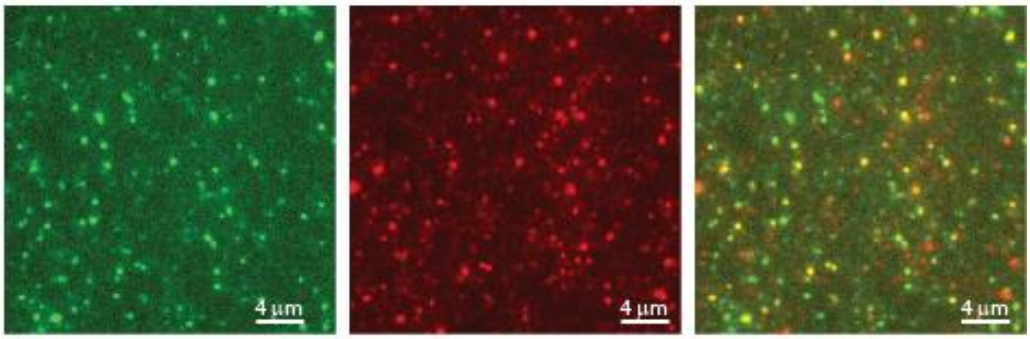
Heteromeric complex stoichiometry measurement in fused cells

To demonstrate the ability to probe heteromeric associations with cell fusion and single molecule imaging, we expressed mNG-ADRβ2 with an mCherry-tagged derivative of ADRβ1 (mCherry-ADRβ1) in a subset of cells. After fusion, both color channels showed distinct puncta and the overlaid image clearly displayed some overlapping spots (Figure 3.6a). The degree of colocalization was quantified by standard pixel level methods (Pearson's and Mander's coefficients) and at the single molecule level by PSF fitting and determining the fraction of spots with a nearest neighbor in the opposite color within 100 nm (Figure 3.6b). After fusion, the respective color channels showed some degree of colocalization, with 26% of mNG-ADRβ2 centroids occurring within 100nm of mCherry-ADRβ2 centroids (Figure 3.6b). This predicts a 40-50% colocalization when corrected for missed colocalized pairs from non-fluorescent proteins. The lower brightness of mCherry prohibited accurate stepwise photobleaching in the red channel; however, based on the single molecule colocalization observed and statistical analysis of heterodimer formation, we conclude that the affinity for ADRβ2 – ADRβ1 heterodimer formation is approximately equal to that of homodimer formation.

CONCLUSION

By achieving detergent-free dilution of protein complexes after physiological assembly, SPReAD facilitates minimally perturbative measurements of subunit stoichiometry for both cytosolic and membrane-bound oligomers. Furthermore, the

a



b

Colocalization Parameter	Mean ±sd (n=7)
Fraction of G/R centroids within 100 nm	0.26 ±0.03
Pearson's coefficient	0.55 ±0.09
Mander's coefficient	0.90 ±0.01
Mander's k1 coefficient	0.64 ±0.31
Mander's k2 coefficient	1.51 ±0.60

Figure 3.6 Two-Color Single Molecule Colocalization Using SPReAD
(a) Green, red and overlay images of mNG-ADRβ2 and mCherry-ADRβ1 revealing colocalization of adrenergic receptor subunits after cell fusion based on PSF colocalization analysis. (b) Averaged colocalization results from single molecule centroid based colocalization analysis and from conventional image-based colocalization methods (n = 7).

use of VSVG as a means of accomplishing cell fusion is highly efficient, nontoxic, and requires only a simple buffer exchange. By contrast, existing methods for probing oligomerization are significantly more complex or disruptive. Use of stimulated emission depletion to reduce excitation volumes by >100-fold can extend the upper limit on FCS measurements, but requires complicated optics and increases photobleaching and phototoxicity. As mentioned earlier, single molecule pull-down approaches are capable of probing molecular heterogeneity in oligomerization, but require extraction of protein complexes from their native environment [4, 26]. Meanwhile, efforts to apply conventional imaging or localization microscopy to stoichiometry analysis rely on *a priori* assumptions about protein distribution or fluorophore blinking [27, 28]. Compared to these methods, SPReAD offers unique advantages, affording single molecule sensitivity for oligomerization studies while maintaining a more physiologically relevant setting.

Although the use of cell fusion for dilution is both simple and rapid, the dynamics of various intracellular processes need to be considered when interpreting results. Syncytia form almost instantly after pH drop, but protein redistribution is diffusion-limited, and thus, much slower for membrane-bound proteins undergoing 2D diffusion compared to cytosolic proteins moving much more rapidly in 3D. This yields two possible modes of analysis: an equilibrium mode, where the final concentration of labeled protein complexes is uniform and proportional to the initial concentration (divided by the co-plating ratio), and a non-equilibrium mode, where concentrations across the imaging dish are non-uniform.

The latter mode, typically carried out 10-45 minutes after buffer exchange for the proteins studied here, is most conducive to measuring subunit stoichiometry because it minimizes the time during which non-covalent protein complexes are able to dissociate. Beyond these considerations of diffusion and dissociation, syncytia appeared to be morphologically stable for 5-6 hours, but it is still largely unclear how the intracellular environment is reshaped during the fusion process. Understanding how syncytium formation affects major signaling pathways will be critical for proper interpretation of SPReAD results.

The brief pH drop required to initiate VSVG-mediated cell fusion may be a potential concern for some applications, since it could alter the stoichiometry of non-covalent membrane complexes exposed to the stimulation buffer, resulting in errors in stoichiometry determination. Higher levels of oligomer formation were consistently observed using SPReAD, in contrast to those observed using the detergent isolated preparation methods required for single molecule pull-down. These results indicate that the brief change in buffer pH is not an issue for the systems we tested. We did find that intracellular pH is largely unaffected, and therefore, proteins and protein complexes en route to the membrane would not be affected. Proteins which form complexes after reaching the membrane (e.g. EGFR) would be able to associate normally during the fusion time course following the pH jump. Recently, a new family of viral fusogens (Reovirus Fast proteins) have been identified [29] which do not require a pH jump to activate. These have potential for use in cell-fusion based single molecule studies in cases where the pH jump is thought to be a problem.

A number of strategies may be used to augment the SPReAD technique and build upon its versatility. The non-equilibrium mode of measurement lends itself to an interesting form of single-cell analysis that can be accomplished by using large co-plating ratios and limiting the time from fusion to measurement/fixation. In this way, proteins from neighboring labeled cells do not immediately mix which may allow for single molecule observations during oligomerization. For true measurement of physiologically relevant interactions, endogenous proteins can be labeled using prevalent genome editing techniques, or ideally, primary cells can be extracted from genetically modified organisms to understand tissue-specific phenotypic variation. Future work may also extend SPReAD applications beyond the cytoplasm and plasma membrane by utilizing membrane contact sites between organelles, examining proteins that exchange between the cytoplasm and other compartments, or by the retargeting of proteins through signal sequence engineering.

By removing limits on expression levels compatible with single molecule experiments without requiring chemical agents for dilution, SPReAD permits minimally perturbative measurements in a variety of cell lines. Aside from the FCS- and stepwise photobleaching-based analyses of subunit stoichiometry highlighted here, we expect that SPReAD will enhance other methods traditionally limited to working at low concentrations such as smFRET, single-particle tracking, and single molecule spectroscopy, thus providing a powerful addition to the single molecule toolkit.

METHODS

Cloning of inducible VSVG and labelled proteins

To avoid the deleterious effects of long-term VSVG expression, the coding sequence for VSVG (Addgene #8454) was cloned into the BamHI and EcoRI sites of the lentiviral pLV Puro Tet vector for doxycycline-inducible expression. A constitutively expressed mNeonGreen lentiviral plasmid was produced by excising mNeonGreen from mNeonGreen-N1 (Allele Biotech) using NheI and NotI and subcloning it into pCDH-puro (System Biosciences). The resulting plasmid has been deposited to Addgene (plasmid #82724).

Synthetic dimers of fluorescent proteins were produced by placing a helical linker $A(EAAAK)_5A$ after the mNG sequence in mNG-C1 (between the BspEI and BglII sites). mNG or mCh2 were then PCR amplified and placed after this linker (between NotI and SpeI sites) to generate mNG-mNG or mNG-mCh2, respectively. The rigid helical linker spaced the fluorescent protein domains further apart to reduce energy transfer [30, 31]. pCDH-puro and mNG-C1 were both digested with NheI and BamHI to excise the fluorescent protein and place it into the pCDH lentiviral plasmid to generate pCDH-puro-mNG, which was used to produce a stable mNG cell line.

mNG-tagged ADRβ2 and EGFR were generated by cloning the fluorescent proteins into the pSNAPf-ADRβ2 backbone (New England Biolabs). mNG was PCRed from mNG-C1 and placed between the EcoRI and SbfI sites of pSNAPf-ADRβ2 (replacing the SNAP tag) to yield mNG-ADRβ2. Site-directed mutagenesis was used to remove a ClaI site from wildtype EGFR. This mutated EGFR was then

PCR amplified and placed between the SbfI and XhoI sites of the pSNAPf-ADRβ2 plasmid, replacing ADRβ2. The EGFR signal sequence was purchased as a gBlock (Integrated DNA Technologies) and placed between the ClaI and BmtI sites to generate mNG-EGFR. Lentiviral versions of mNG-ADRβ2, mNG-EGFR, mNG-mGluR3, and mNG-mGluR5 were produced by amplifying each plasmid via PCR and digesting with XbaI and NotI to place the fusion protein after the CMV promoter in pCDH-puro. To make Orai1-mNG, Orai1-YFP (Addgene #19756) and mNG-N1 were digested with AgeI and NotI to remove YFP and replace it with mNG.

Cell culture and generation of stable cell lines

U2OS human osteosarcoma cells were cultured in DMEM without phenol red, supplemented with 10% FBS, sodium pyruvate, 1x Glutamax, and 1x antibiotic-antimycotic; all cell culture media and supplements were purchased from Life Technologies. For stable expression of VSVG under tetracycline control, U2OS cells were first stably transduced with the rtTA NeoR plasmid for the reverse tetracycline-controlled transactivator (rtTA) protein. Lentiviral particles were generated in HEK293 cells and used to transduce U2OS cells as previously described[32]. Stably transduced cells were selected using 700 µg/mL G418. U2OS rtTA cells were then transduced with pLV puro Tet-VSVG and selected using 2 µg/mL puromycin. Doxycycline was withheld from cell culture media until 24 hours prior to cell fusion. Stable mNeonGreen cell lines were produced by transducing U2OS Tet-VSVG cells with pCDH-puro-mNG-C1, pCDH-puro-mNG-ADRβ2, and pCDH-puro-mNG-EGFR, followed by an antibiotic selection using 2 µg/mL puromycin. Stable mNeonGreen-mGluR expressing HEK293T cell lines were

produced by transducing cells with either pCDH-puro-mNG-mGluR3 or pCDH-puro-mNG-mGluR5 and enriched by fluorescence-activated cell sorting. (BD Biosciences FACSAria).

Fusion Assay

U2OS Tet-VSVG cells were plated onto collagen coated glass-bottom dishes. After reaching confluence, fresh media with 2 µg/mL doxycycline was added and the cells were returned to a CO_2 incubator for 24 hours. Cells were then fused by removing culture media, washing with PBS, and incubating in fusion buffer (PBS with 25 mM MES, pH 5.5) for 1 to 5 minutes. Cells were washed with PBS and fresh culture media was added before returning cells to the CO_2 incubator. Cell membranes and nuclei were labelled at various time points by incubating with 5 µg/mL Wheat Germ Agglutinin Alexa 647 (Life Technologies) and 5 µg/mL Hoechst in Hank's balanced salt solution for 10 minutes prior to fixation with 4% paraformaldehyde. Fixed cells were imaged on a spinning disk confocal microscope (Olympus) with air objectives (40x/0.9, 20x/0.7 and 10x/0.4) and examined for syncytia formation.

Confocal Microscopy and Fluorescence Correlation Spectroscopy

U2OS Tet-VSVG cells were transfected with FP control plasmids or FP-tagged PKA-subunits using Lipofectamine 3000; for cytoplasmic mNG measurements, stable U2OS mNG cells were used to accurately control the number of expressing cells. Serum-free Fluorobrite DMEM (Life Technologies) was used to minimize cellular autofluorescence. The two were mixed at various ratios and $5x10^5$ cells were plated in the well of a 14 mm diameter glass-bottom

dish (collagen/fibronectin-coated) using doxycycline-supplemented media (2 µg/mL); additional media was added 2-12 hours after plating, after cells were visibly attached and spread. Cells were imaged on a confocal microscope (Zeiss LSM880). Fluorescence correlation spectroscopy was performed on the same instrument using the LSM880 32-channel GaAsP detector in photon counting mode. Standard FCS fitting equations were used (10). For non-PKA FCS measurements, the data was fit to a single component diffusion with triplet model. Absolute concentrations for cytoplasmic mNeonGreen were obtained by calibrating the focal volume with known concentrations of Alexa 488. From the two-color cross-correlation measurements, the average number of particles was determined using:

$$N_{G,R} = \frac{1}{G(0)_{G,R}} \qquad \frac{N_x}{N_G} = \frac{G(0)_x}{G(0)_R} \qquad \frac{N_x}{N_R} = \frac{G(0)_x}{G(0)_G}$$

where $N_{G,R}$ is the number of green or red particles, and N_X/N_G and N_X/N_R are the heterodimer fractions. For Protein Kinase A experiments, PKA-transfected U2OS cells were mixed 1:10 with non-expressing VSVG cells and incubated in doxycycline-supplemented Fluorobrite DMEM for 24 hours. Cells were then fused by a 5-minute incubation in fusion buffer and FCS was performed in syncytia one hour later. In order to maintain the same syncytial position for post-stimulation measurements, 2x cAMP-stim buffer (50 µM forskolin, 200 µM IBMX in Fluorobrite DMEM) was added directly to the imaging dish in equal volume to the residual media and a second FCS recording was initiated 5 minutes later. PKA data was fit to a two-component diffusion model [10].

Single-Molecule Imaging After Cell Fusion

U2OS Tet-VSVG cells were transfected with FP-tagged receptor constructs and plated onto glass-bottom dishes with non-transfected cells at a ratio of 10:1 (non-transfected:transfected), as described above. After 24h of doxycycline induction, cells were fused and imaged live (1-2 hours after fusion) or fixed for stoichiometry/colocalization analysis. Cells were fixed with 4% paraformaldehyde for 3 hours in the dark at room temperature to eliminate residual mobility of membrane proteins after short fixation[33]. For mNG-EGFR experiments, the syncytia were stimulated with 200 ng/mL EGF 75 minutes after cell fusion was initiated, and fixed 5 minutes later, or fixed without EGF treatment.

TIRF Microscopy

A custom-built azimuthal-scanning objective-TIRF microscope was used for single molecule imaging. Excitation at 488 nm and 561 nm were used to excite mNeonGreen and mCherry, respectively, and were directed to the sample using a quad polychroic (ZT405/488/561/640rpc, Chroma Technology) housed in the filter wheel. A beam telescope and focusing lens were used to create a collimated beam out of the objective (Olympus UApoN 100x/1.49), while a pair of galvanometer mirrors (Cambridge Technology) controlled the angle of incidence. For the detection path, a TuCam adaptor (Andor) equipped with band pass filters (ET525/50 for mNeonGreen and ET605/52m for tdTomato) was used to split emissions onto two EMCCDs (Andor iXon 887 and 897 Ultra). Image coregistration was accomplished by acquiring brightfield images of a calibration objective (Zeiss calibration LSM) prior to each imaging experiment and ensuring that the images

were coregistered to better than one pixel over the camera field-of-view through alignment of the detection pathway. Live-cell data was acquired at 37°C using an objective heater (Bioptechs), while fixed cell experiments were performed at room temperature. Coverslips were scanned for regions with a suitable density of molecules for single molecule analysis; regions with unfused fluorescent cells or too few/too many molecules were avoided. For bleach step analysis, 2000 frames were recorded at 10-30 Hz; laser intensity was kept low to mitigate blinking artifacts. For colocalization analysis, 20 frames were acquired and averaged during post-processing.

Single Molecule Data Analysis

Photobleaching movies were analyzed using three different methods. Two methods are based on the use of a custom lab software package (ImageC.exe, written in C/C++ under Microsoft Visual Studio 2017) and the third involved the use of a published automated software package written in Matlab called Progressive Image Filtering (PIF). [32]. In both programs, molecules (PSFs) were first located automatically by successive processing of the summed image stack to locate fluorescent puncta above a certain threshold that meets a specified Gaussian fit criterion. For each molecule, an ROI (typically 5x5) centered around the pixel containing the PSF centroid was created and the ROI mean values vs. time (frame) extracted from the stack. The ROI center pixel coordinate was readjusted slightly as needed as the data is extracted from the frames so that the brightest pixel is always at the center. ROI fluorescence traces of all the spots located are stored within the program and displayed as time trace plots for manual (i.e. visual) step

counting in ImageC, or used for with the automated step-finding algorithms in ImageC or PIF. Both algorithms count the number of bleach steps based on signal noise and a user-set change in the trace count level that determines a valid step. Traces without discernible bleach steps were discarded. At least 700 molecules were analyzed for each sample. Further information on the programs used is provided in Appendix B.

For colocalization analysis, data from two EMCCDs were analyzed to find spots in both the green and red channels using either a custom MATLAB script or function built into our lab's custom analysis program (ImageC). The PSFs were fit to a Gaussian model to determine center locations. A colocalization fraction was calculated to be the fraction of mNeonGreen spots with an mCherry spot less than 100 nm away.

Substrate Preparation

To minimize glass autofluorescence and maximize cell attachment, plain glass-bottom dishes were cleaned and coated with fibronectin. Dishes were etched with 1 M KOH for 20 minutes, followed by a water and then PBS rinse. For fibronectin coating, dishes were incubated in 4% (3-Mercaptopropyl)trimethoxysilane (Sigma-Aldrich) in ethanol for 30 minutes, rinsed with ethanol, incubated with (N-γ-maleimidobutyryl-oxysuccinimide ester) crosslinker (4 mM in ethanol, Thermo Scientific), rinsed with ethanol and dried thoroughly in a sterile biosafety cabinet. Dishes were then incubated with 5 µg/mL fibronectin for 2 hours at room temperature, followed by overnight at 4°C, then rinsed with PBS and stored in PBS at 4°C until use (up to several weeks).

REFERENCES

1. T. Berggard, S. Linse, P. James, Methods for the detection and analysis of protein-protein interactions. *Proteomics* **7**, 2833-2842 (2007).

2. B. N. Giepmans, S. R. Adams, M. H. Ellisman, R. Y. Tsien, The fluorescent toolbox for assessing protein location and function. Science 312, 217-224 (2006).

3. M. H. Ulbrich, E. Y. Isacoff, Subunit counting in membrane-bound proteins. Nat Methods 4, 319-321 (2007).4. Jain, A. *et al.* Probing cellular protein complexes using single-molecule pull-down. *Nature* **473,** 484–8 (2011).

5. S. Kalipatnapu, A. Chattopadhyay, Membrane protein solubilization: recent advances and challenges in solubilization of serotonin1A receptors. IUBMB Life 57, 505-512 (2005).

6. A. Gottesman, J. Milazzo, Y. Lazebnik, V-fusion: a convenient, nontoxic method for cell fusion. Biotechniques 49, 747-750 (2010)..

7. B. M. Ogle, M. Cascalho, J. L. Platt, Biological implications of cell fusion. Nat Rev Mol Cell Biol 6, 567-575 (2005).).

8. D. Duelli, Y. Lazebnik, Cell-to-cell fusion as a link between viruses and cancer. Nat Rev Cancer 7, 968-976 (2007).

9. D. Madge, E. Elson, W. W. Webb, Thermodynamic Fluctuations in a Reacting System---Measurement by Fluorescence Correlation Spectroscopy. Physical Review Letters 29, 705 (1972).

10. K. Bacia, S. A. Kim, P. Schwille, Fluorescence cross-correlation

spectroscopy in living cells. Nat Methods 3, 83-89 (2006).

11. J. Mutze, T. Ohrt, P. Schwille, Fluorescence correlation spectroscopy in vivo. Laser Photonics Rev 5, 52–67 (2011).

12. T. J. Stasevich et al., Cross-validating FRAP and FCS to quantify the impact of photobleaching on in vivo binding estimates. Biophys J 99, 3093-3101 (2010).

13. Chen, Y., Müller, J. D., Ruan, Q. & Gratton, E. Molecular brightness characterization of EGFP in vivo by fluorescence fluctuation spectroscopy. Biophys. J. 82, 133–44 (2002).

14. H. Park, C. Pack, M. Kinjo, B. K. Kaang, In vivo quantitative analysis of PKA subunit interaction and cAMP level by dual color fluorescence cross correlation spectroscopy. Mol Cells 26, 87-92 (2008).

15. H. Yin, A. D. Flynn, Drugging Membrane Protein Interactions. Annu Rev Biomed Eng 18, 51-76 (2016).16.

16. R. J. Arant, M. H. Ulbrich, Deciphering the subunit composition of multimeric proteins by counting photobleaching steps. Chemphyschem 15, 600-605 (2014).17.

17. I. Bang, H. J. Choi, Structural features of beta2 adrenergic receptor: crystal structures and beyond. Mol Cells 38, 105-111 (2015).

18. C. Romano, W.L. Yang, K. L. O'Malley, Metabotropic glutamate receptor 5 is a disulfide-linked dimer. *J Biol Chem* **271**, 28612-28616 (1996).

19. A. Penna et al., The CRAC channel consists of a tetramer formed by Stim-

induced dimerization of Orai dimers. Nature 456, 116-120 (2008).

20. X. Cai et al., The Orai1 Store-operated Calcium Channel Functions as a Hexamer. J Biol Chem 291, 25764-25775 (2016).

21. W. Ji et al., Functional stoichiometry of the unitary calcium-release-activated calcium channel. Proc Natl Acad Sci U S A 105, 13668-13673 (2008).20

22. A. Gschwind, O.M. Fischer, A. Ullrich, The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat. Rev. Cancer* 4**,** 361–70 (2004).

23. I. Chung et al., Spatial control of EGF receptor activation by reversible dimerization on living cells. Nature 464, 783-787 (2010)..

24. X. Yu, K. D. Sharma, T. Takahashi, R. Iwamoto, E. Mekada, Ligand-independent dimer formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced EGFR signaling. Mol Biol Cell 13, 2547-2557 (2002).

25. A. H. Clayton, S. G. Orchard, E. C. Nice, R. G. Posner, A. W. Burgess, Predominance of activated EGFR higher-order oligomers on the cell surface. Growth Factors 26, 316-324 (2008).

26. H. W. Lee et al., Real-time single-molecule co-immunoprecipitation analyses reveal cancer-specific Ras signalling dynamics. Nat Commun 4, 1505 (2013)..

27. A. G. Godin et al., Revealing protein oligomerization and densities in situ using spatial intensity distribution analysis. Proc Natl Acad Sci U S A 108, 7010-7015 (2011).

28. F. Fricke, J. Beaudouin, R. Eils, M. Heilemann, One, two or three? Probing the stoichiometry of membrane proteins by single-molecule localization microscopy. Sci Rep 5, 14072 (2015).

29. M. Ciechonska, R. Duncan, Reovirus FAST proteins: virus-encoded cellular fusogens. Trends Microbiol 22, 715-724 (2014).

30. R. Arai, H. Ueda, A. Kitayama, N. Kamiya, T. Nagamune, Design of the linkers which effectively separate domains of a bifunctional fusion protein. Protein Eng 14, 529-532 (2001).

31. X. Chen, J. L. Zaro, W. C. Shen, Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev 65, 1357-1369 (2013).

32. H. McGuire, M. R. Aurousseau, D. Bowie, R. Blunck, Automating single subunit counting of membrane proteins in mammalian cells. J Biol Chem 287, 35912-35921 (2012).

CHAPTER 4

INVESTIGATING THE STOICHIOMETRY OF METABOTROPIC GLUTAMATE RECEPTORS *IN VIVO* USING A NOVEL SINGLE MOLECULE IMAGING TECHNIQUE

Metabotropic glutamate receptors (mGluRs) play an important role in neuromodulation and in pathologies including schizophrenia and Fragile X syndrome. Functional homodimers have been observed in all mGluRs and heterodimers have been observed in groups II and III. However, interactions between the members of group I have not yet been thoroughly investigated. We sought to elucidate the stoichiometry of these complexes, but did not find significant evidence of heterodimer formation between any pair of mGluRs. We verified our findings using stepwise photobleaching, fluorescence colocalization, and fluorescence lifetime imaging microscopy.

INTRODUCTION

Glutamate acts as both a neurotransmitter and neuromodulator in the central nervous system. The excitatory effects are mediated by ionotropic glutamate receptors (iGluRs) that function as ion channels. The neuromodulatory roles are mediated by metabotropic glutamate receptors (mGluRs) acting via secondary messengers [1].

Metabotropic glutamate receptors (mGluRs) are members of a family of class C G-protein-coupled receptor (GPCR) which are widespread throughout the central nervous system (Figure 4.1a) [1]. They play a role in many processes such as learning and memory, and their malfunction has been linked to anxiety, schizophrenia, and neurotoxicity [6-13]. In particular, an increase in mGluR5

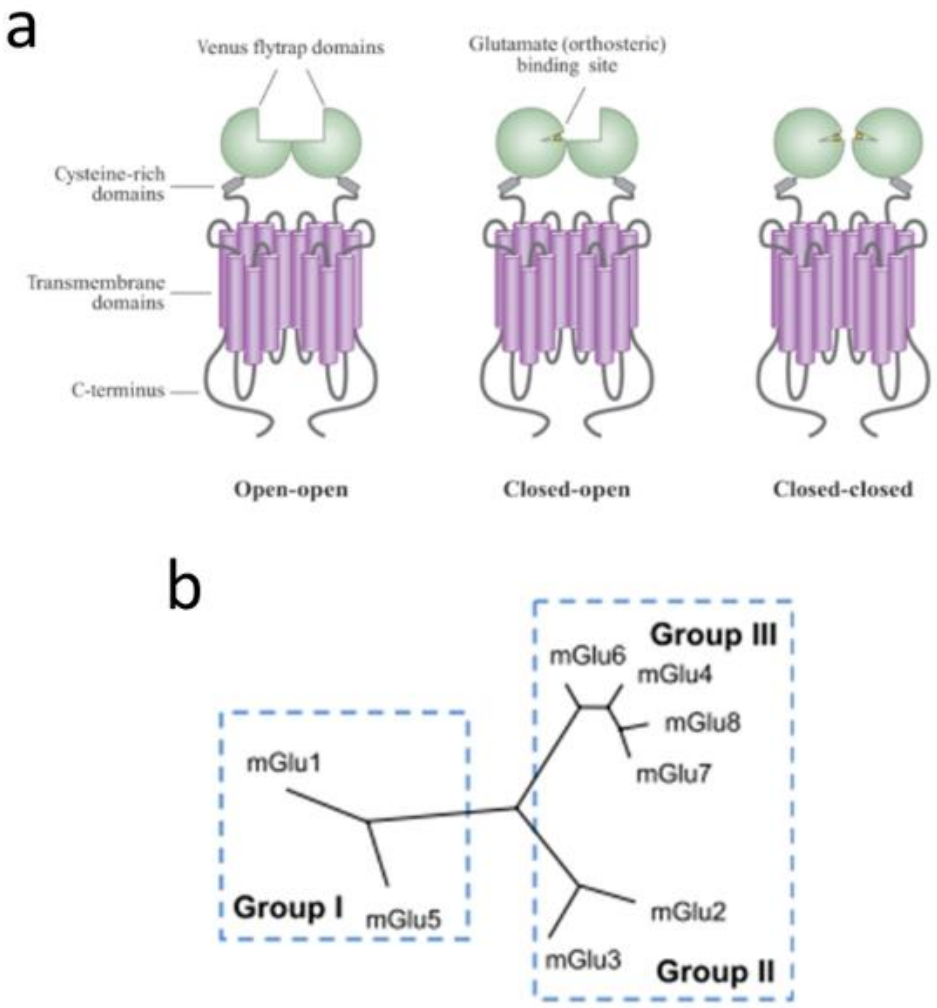


Figure 4.1 Metabotropic Glutamate Receptors and Their Classification
(a) Diagram of metabotropic glutamate receptors showing extracellular ligand binding domain states with and without ligand bound (adapted from *Niswender et al* [2]). (b) The categorization of mGluRs. Areas surrounded in blue boxes have some evidence showing interactions between the mGluRs contained within (adapted from *Doumazane et al* [3]).

signaling has been linked to the cognitive symptoms accompanying fragile X syndrome [14, 15].

Eight mGluRs (mGluR1-8) have been identified and divided into three groups (I-III) based on signal transduction pathway, agonist selectivity, and sequence homology (Table 4.1, Figure 4.1b). These characteristics have been elucidated using structural, biochemical, optical, and physiological methods [3, 4, 16-20]. The physiological roles of each mGluR are difficult to generalize because their function is highly dependent on neuronal population and location relative to the synapse [2].

It has long been known that one functional form of mGluRs is that of a stable homodimer [4, 17, 18]. This structure is stabilized by two features on the extracellular domain: hydrophobic regions on the dimer interface and a disulfide bridge between Cys residues formed in the endoplasmic reticulum prior to trafficking (Figure 4.2a) [18-20]. A number of studies over the last decade have demonstrated that functional heterodimers exist between members of group II and III, with strong evidence suggesting that mGluR2 and mGluR4 preferentially form heterodimers (Figure 4.2b) [3, 5, 20, 21]. The interactions between group I mGluRs are not as well understood as those of group II and III. Co-immunoprecipitation studies demonstrate that group 1 mGluRs do not form heterodimers, but subsequent FRET evidence has shown that they do associate in some manner [3, 4]. Further functional studies found that there is an interaction between mGluR1 and mGluR5 but that it does not appear to be that of a heterodimer (Figure 4.2c both panels) [22].

Key features of mGluRs

Group	Receptor/splice variants	CNS expression	Synaptic localization	Signaling pathways of group
Group I	mGluR1 a,b,c,d,e,f Taste mGluR1	Widespread in neurons Taste buds	Predominantly postsynaptic	Phospholipase C stimulation Stimulation of adenylyl cyclase (some systems) MAP kinase phosphorylation
	mGluR5 a,b	Widespread in neurons, astrocytes		
Group II	mGluR2	Widespread in neurons	Presynaptic and postynaptic	Inhibition of adenylyl cylcase Activation of K^+ channels Inhibition of Ca^{++} channels
	mGluR3 GRM3A2 GRM3A4 GRM3A2A3	Widespread in neurons, astrocytes		
Group III	mGluR4 Taste mGluR4	Widespread in neurons, High in cerebellum Taste buds	Predominantly presynaptic	Inhibition of adenylyl cylcase Activation of K^+ channels Inhibition of Ca^{++} channels
	mGluR6 a,b,c	Retina	Postsynaptic in ON-bipolar retinal cells	Stimulation of cGMP phosphodiesterase (mGluR6)
	mGluR7 a,b,c,d,e	Widespread in neurons	Active zone of presynaptic terminals	
	mGluR8 a,b,c	Lower and more restricted expression than mGluR4/7	Predominantly presynaptic	

Table 4.1 Key Features of mGluRs

A table listing metabotropic glutamate receptor (mGluR) group categorization, location in central nervous system, location relative to synapse, and the signaling pathways used (adapted from *Niswender et al* [2]

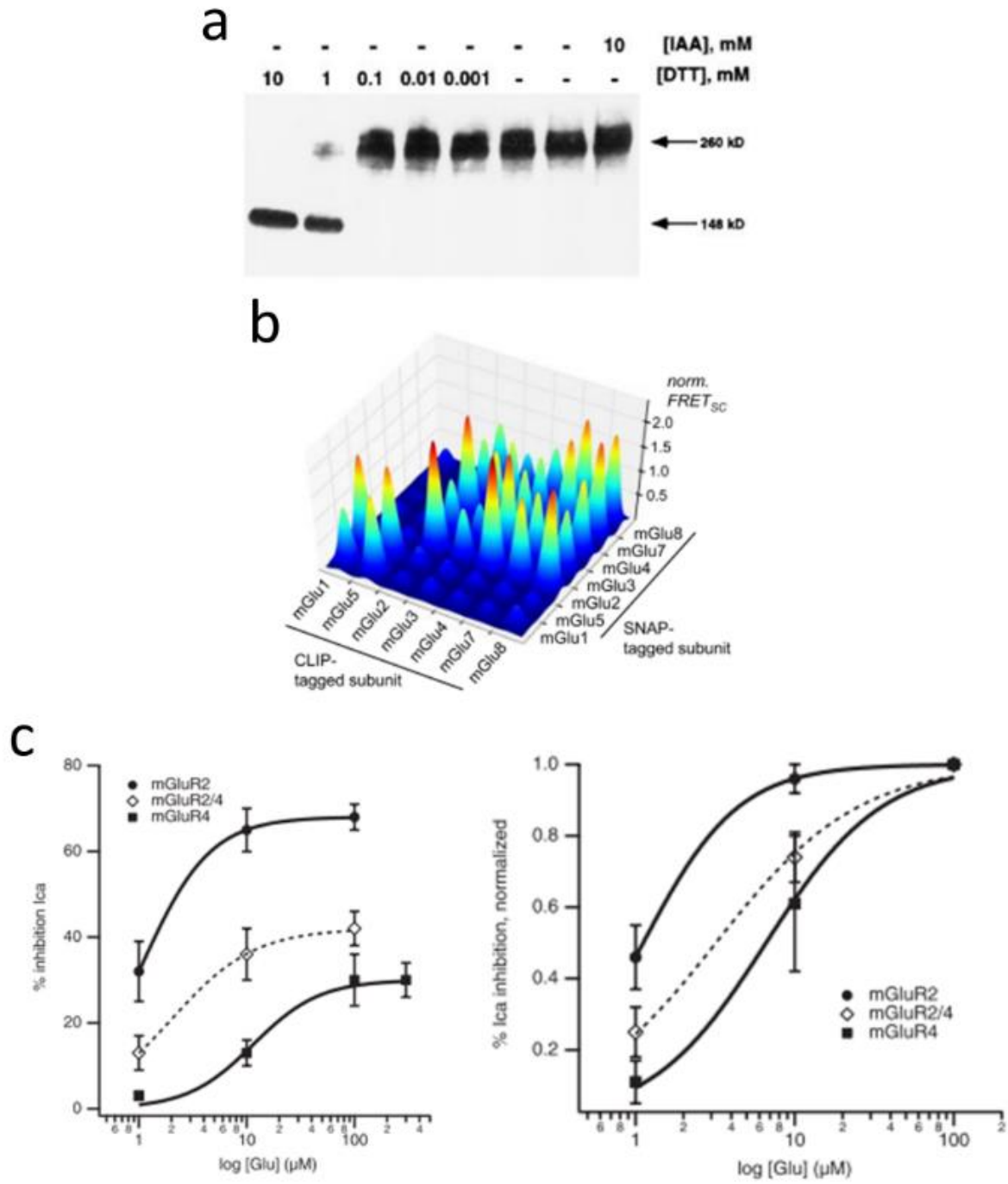


Figure 4.2 Architecture of Metabotropic Glutamate Receptor Complexes
(a) Western blot analysis using an antibody directed at mGluR5 showing a shift to lower mass in the presence of DTT (adapted from *Romano et al* [4]). (b) FRET intensity of mGluRs combinations labeled with CLIP and SNAP attached FRET donor and acceptor molecules. Interactions within group I and between the members of groups II and III are demonstrated (adapted from *Doumazane et al* [3]). (c) Glutamate dose-response curves showing that receptor complexes formed when mGluR2 and mGluR4 are co-expressed in a cell respond to stimulation differently than those formed by mGluR2 or mGluR4 alone (adapted from *Kammermeier* [5]).

primary pharmacological method used to alter the sensitivity of a specific receptor to glutamate stimulation [2]. Determining which mGluRs are able to interact with one another is an important consideration in the development of therapeutic agents, as these allosteric modulators are specifically developed to effect only a single type of mGluR. It has been demonstrated that a NAM must occupy allosteric sites in both subunits of a receptor complex to effectively block signaling and therefore it was postulated that heterodimers would need two unique NAMs to effectively silence the receptor [23]. This was confirmed when functional studies of mGluR2-mGluR4 heterodimers showed that signaling was only repressed when two different NAMs were used, one specific to each subunit [5].

We intend to investigate the interaction between the group I mGluRs, mGluR1 and mGluR5, using stepwise photobleaching, fluorescence colocalization fluorescence lifetime imaging microscopy (FLIM), and two-color fluorescence cross-correlation (FCCS). Protein complexes will be examined both *in vitro* and *in vivo* using single molecule pull-down (SiM-Pull) and single protein recovery after dilution (SPReAD) respectively.

RESULTS

Determination of best cell line to express mGluRs for imaging

Given that native mGluRs function in neurons, our first course of action was to determine which of our cell lines best preserved the formation of the well-studied homodimers. Stable U2OS and HEK293T cell lines expressing a single mNeonGreen (mNG) labeled mGluR1a were created and stepwise photobleaching data was collected to determine the fraction of receptor complexes which

photobleached in two steps, an indication of dimeric stoichiometry (Figure 4.3a). The functional form of singly expressed mGluRs is that of a homodimer, therefore a higher two-step bleach fraction was interpreted as evidence that a cell line was more suitable to use as an expression system. Complexes expressed in HEK293T cells displayed 4% more two-step bleaches and 4% less one-step bleaches than those expressed in U2OS cells (Figure 4.3b). As a result of this experiment, HEK293T cells were determined to be the more favorable expression system. It is worth mentioning that these results were similar to, but slightly less than, previously published values [20].

Receptor complex stoichiometry in cells expressing a single mGluR

Stable HEK293T cell lines expressing a single mNG tagged mGluR were imaged in order to count the number of subunits present in the receptor complexes. This was accomplished by observing the stepwise photobleaching of mNG. Each mGluR was found to have a ~33% two-step bleach percentage (Figure 4.3c). The number of subunits can be calculated from this data by correcting for the percent of successfully matured fluorophores, which we set at 80%. mGluR1a, 3, and 5 were calculated to have a 33% dimer percentage and mGluR4 was calculated to have a 37% dimer percentage. These measurements also provided a large data set that was manually counted and used to calibrate parameters on automated stepwise photobleaching analysis algorithms (see Appendix 2 for more detail).

The effect of a reducing agent on mGluR complex stoichiometry

The role of the cysteine bridge between subunits in the receptor complex

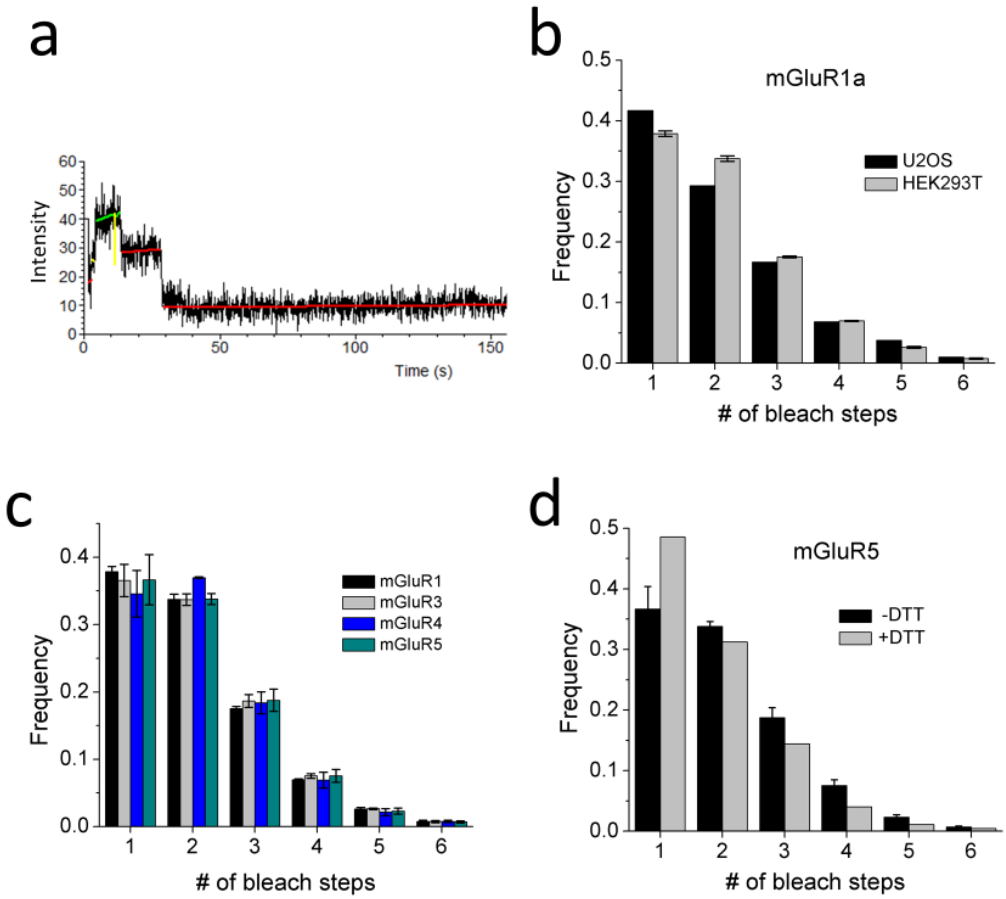


Figure 4.3 Single Color mGluR Complex Stoichiometry Measurements
(a) Example photobleaching trace showing two step bleaching. (b) Comparison of the bleach step counts of mNG-mGluR1a expressed in U2OS and HEK293T cells (mean ± SEM, n = 3 for HEK293T cells). (c) Bleach step histograms for mGluR1,3,4,5 (mean ± SEM, n = 3) (d) Comparison of bleach step distribution between mNG-mGluR5 without special preparation and with a 15 minute incubation in 15 mM DTT (mean ± SEM, n = 3 for -DTT sample).

was examined by gathering stepwise photobleaching data from cells treated with 15 mM dithiothreitol (DTT) during the last 15 minutes of cell fusion. We found no significant decrease in the two-step bleach fraction in the presence of DTT (Figure 4.3d). This finding suggests that the cysteine bridge between subunits is not forming in our cells and that the dimer complex is stabilized solely by hydrophobic interactions on the dimer interface. Further experimentation with longer DTT incubation times would clarify whether the cysteine bridges are not forming or that 15 minutes is not long enough to break a large fraction of the bonds.

Fluorescence colocalization of cells expressing pairs of mGluRs

In order to investigate the heteromeric interactions between different mGluRs, we created cell lines expressing two mGluRs, one tagged with mNG and the other with mScarlet (mSc) (Figure 4.4a). We then attempted to establish a standard for colocalization fractions by measuring the fluorescence colocalization of two positive control complexes, mNG-mGluR1a + mSc-mGluR1a and mNG-mGluR2 + mSc-mGluR4, and two negative control complexes, mNG-mGluR1a + mSc-mGluR4 and mNG-mGluR5 + mSc-mGluR4 (Figure 4.4b). These positive controls were selected because mGluR2 and mGluR4 preferentially form heterodimers and the formation of homodimers has already been established [3-5, 20]. The negative controls were specifically chosen because it has been demonstrated that group I mGluRs do not interact with the other groups [3]. A simulated data set of randomly distributed spots was also generated to find a baseline for coincidental colocalization (Figure 4.4b, rightmost bar).

Cells expressing mGluR1a and mGluR5 were imaged and the frequency of

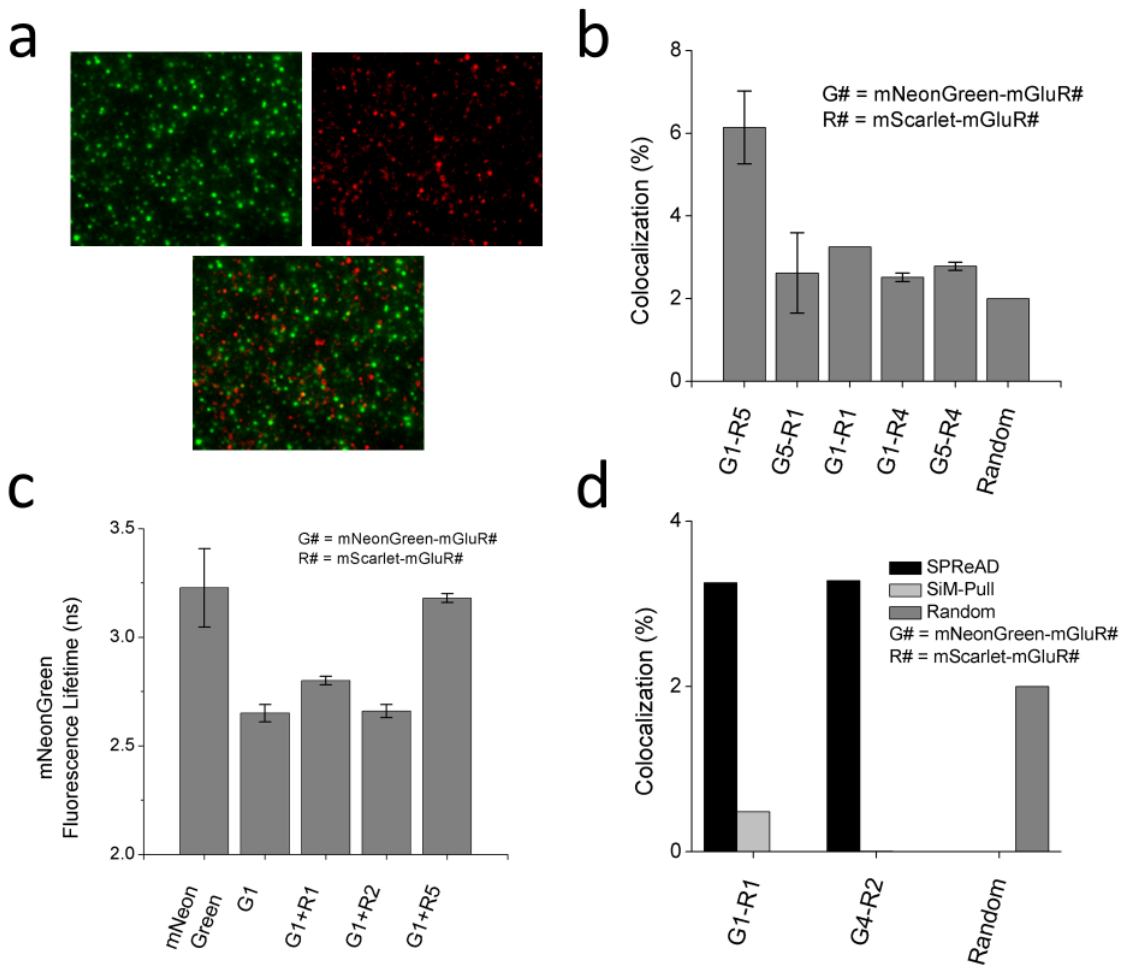


Figure 4.4 Fluorescence Colocalization and Lifetime Imaging of mGluR Complexes (a) Example of colocalization images with mNeonGreen-mGluR1a (upper left), mScarlet-mGluR4 (upper right), and a merged image (lower) (b) Colocalization fraction between mNG and mSc labeled mGluR pairs. “Random” is the colocalization between two 512x512 images containing 600 spots, which is approximately the same number as the mGluR data sets (mean ± SEM, n = 2). (c) Comparison of the fluorescence lifetime of mNG in solution and as part of several fusion protein complexes. (d) (c) Comparison of colocalization fraction between SPReAD and SiM-Pull (mean ± SEM, n = 10).

colocalization was measured to determine if they formed heterodimers. These measurements were taken using both combinations of fluorophore-receptor fusion proteins, mNG-mGluR1a + mSc-mGluR5, and mSc-mGluR1a and mNG-mGluR5, to investigate whether or not there was an observable interaction between a specific fluorophore and the labeled receptor. Furthermore, this enabled subunit counting of each protein, as mSc is not suitable for stepwise photobleaching.

We found the rates of colocalization between the negative controls, positive controls, and the complexes of interest to be similar to one another, and close to the colocalization fraction of images with randomly distributed spots. While this degree of colocalization was expected for the negative controls, the degree of colocalization was expected to be much higher for the positive controls.

Fluorescence lifetime imaging to assess mGluR heterodimer formation via Förster resonance energy transfer

We measured the fluorescence lifetime of mNG in the cell lines used for colocalization to ensure our results were consistent between optical methods (Figure 4.4c). If mNG and mSc labeled proteins were in a complex then FRET would occur and their fluorescence lifetime would decrease as the excited electrons in mNG would lose energy to both fluorescence and FRET. Our measured value for the fluorescence lifetime of purified mNG was 3.23 ± 0.18 ns which agrees with literature values [24]. The fluorescence lifetime of mNG in cells expressing only mNG-mGluR1a was 2.65 ± 0.04 ns. This result may be due to self-quenching (FRET) between the mNG labels on each subunit. The lifetime

measured for the positive control, mNG-mGluR1a + mSc-mGluR1a was 2.80 ± 0.02 ns and the negative control, mNG-mGluR1a + mSc-mGluR4 was 2.66 ± 0.03 ns, while the lifetime for mNG-mGluR1a + mSc-mGluR5 was 3.2 ± 0.02 ns. The agreement between mGluR1a and the negative control is expected, as mNG-mGluR1a should strictly be forming homodimers in each case. A possible explanation for the slight increase in lifetime observed in the positive control is that heterodimers are forming and there is some Förster resonance energy transfer (FRET) occurring between mNG and mSc. However, the lifetime we measured in the mGluR1-mGluR5 cells agreed with the lifetime of purified mNG, suggesting that no mGluR1a homodimers are forming and there is no FRET occurring between the fluorophores.

Effects of sample preparation on receptor complex stoichiometry

In order to verify that the SPReAD method of sample preparation was not causing complex dissociation due to the brief drop in pH or during the ~60 minute fusion period, we measured the stoichiometry of complexes in samples prepared using SiM-Pull, which has been used previously in the literature (Figure 4.4d) [20].

We observed that the two-step bleach fraction in single color mGluR expressing HEK293T cells was slightly lower in SiM-Pull prepared cells than in SPReAD cells, supporting our claim that SPReAD is less disruptive to complex stoichiometry (see Chapter 3). The positive control heterodimer experiments were performed using SiM-Pull with an antibody against mNG. These showed minimal mSc pull-down and a similar fraction of colocalization to the fraction observed in SPReAD cells.

Fluorescence cross-correlation spectroscopy to investigate heterodimerization

In order to further investigate these heterodimeric interactions, we performed fluorescence cross correlation spectroscopy (FCCS) on SPReAD cells expressing mNG-mGluR1 and mSc-mGluR5 (Figure 4.5a). We measured no cross correlation in each of the two color samples, eliminating our theory that mGluR1-mGluR5 were preferentially forming heterodimers (Figure 4.5b).

CONCLUSION

We found that mGluRs labeled with an N-terminal mScarlet (mSc) were unable to form dimers with N-terminally mNeonGreen (mNG) labeled mGluRs. Since mSc is a poor candidate for stepwise photobleaching, we were unable to determine the stoichiometry of mSc labeled homomers. Further experiments are necessary to elucidate whether the location of mSc in the fusion protein is the issue or if mSc interacts with mGluRs in a detrimental way regardless of the location. Another more general concern for heterodimer formation in transduced or transfected cells is the preservation of the numerous post-translational modifications that mGluRs undergo prior to membrane trafficking in a physiological setting [25]. One could imagine that the two different mGluR subtypes would need to be spatially coordinated in some specific way to be optimally post-processed to form oligomers. Given the complex structure of the nucleus and the interactions with the ER, randomly inserted genes may not be sufficient to recapitulate the biology.

A published FRET study was performed using mGluRs labeled with Green

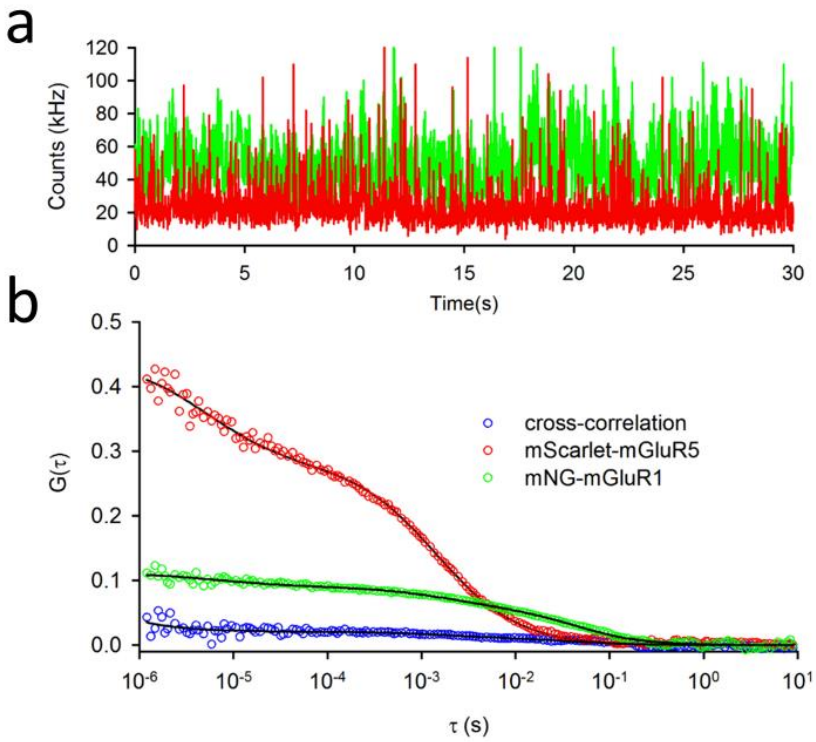


Figure 4.5 Two-color fluorescence cross-correlation spectroscopy of mGluRs
mNG-mGluR1 and mSc-mGluR5 were co-expressed and HEK cells and a SPReAD sample was prepared for imaging. (a) Plot of the counts per second gathered for mNG-mGluR1 and mSc-mGluR5. FCS was possible with these cells due to dilution via the SPReAD technique. (b) Cross-correlation plot showing no cross-correlation between mNG-mGluR1 and mSc-mGluR5

and Lumi4 (donor and acceptor) on their N-terminus [3]. Previous colocalization experiments employed pairs of mGluRs labeled with a C-terminal GFP or mCherry and stepwise photobleaching competition experiments were conducted between a C-terminally GFP labeled mGluR and an unlabeled mGluR [20].

Our plan is to attempt to reproduce the SiM-Pull colocalization experiments using mGluRs labeled with a C-terminal mSc in order to separate mNG and mSc on different sides of the plasma membrane in an attempt to eliminate any interaction between the fluorophores.

METHODS

Cell culture and generation of stable cell lines

Fluorescent mGluRs were produced by placing the signal sequence for mGluR2 at the N-terminus attached to mNG or mSc connected by a 25 AA linker to the rest of the receptor. The mNG variants were then inserted into pCDH-CMV-puro and pCDH-CMV-blast ((Addgene #72265) with puromycin or blasticin resistance inserted) respectively to generate lentivral plasmids used to create stable HEK293T and U2OS mGluR expressing cell lines. Stable cell lines containing only the mNG variant of the mGluRs were generated first and enriched by fluorescence-activated cell sorting. (BD Biosciences FACSAria courtesy of the Cornell BRC). For the cell lines expressing two colors, subsequent transduction of the mNG-mGluR lines with mSc-mGluRs was performed and again enriched using fluorescence-activated cell sorting.

Single Protein Recovery after Dilution

A detailed explanation of Single Protein Recovery after Dilution can be

found in Chapter 3. Cells expressing mGluRs were co-plated with U2OS Tet-VSVG cells onto fibronectin coated glass-bottom dishes at a ratio of approximately 1:10 (mGluR:VSVG). After allowing the cells to attach to the surface for approximately 60 minutes, fresh media with 2 μg/mL doxycyline was added and the cells were returned to a CO_2 incubator for 24-36 hours. Cells were then fused by removing culture media and incubating in fusion buffer (PBS with 25 mM MES, pH 5.5) for 3 minutes. The fusion buffer was removed and culture media was restored before returning cells to the CO_2 incubator for 60-70 minutes. Cells were then rinsed with PBS and incubated with 4% paraformaldeyde (Electron Microscopy Sciences) in PBS for 4 hours in the dark at room temperature. PFA was removed and the cells were rinsed again using PBS and then stored in the dark at 4°C for no longer than 36 hours prior to imaging.

Single Molecule Pull-down

SiM-Pull flow chambers were prepared using a modified version of the protocol used in [26]. Glass slides and coverslips were prepared for surface passivation using successive 10 minute sonication periods in 2% Alconox (Alconox), acetone, and 1 M NaOH with a thorough milliQ water rinse in between each step. After a final rinse with milliQ water, the glass was blow-dried using a stream of nitrogen. The surface was functionalized using n-(2-aminoethyl)-3-aminopropyltrimethoxysilane (UCT Specialties) diluted in an acidic methanol solution (5% acetic acid) to a final volumetric ratio of 1:100. The slides and coverslips were incubated in this solution for 10 minutes prior to a 1 minute sonication and final 10 minute incubation. Slides and coverslips were then rinsed

in milliQ and blow-dried as above. The PEGylation was performed by adding a 1M sodium bicarbonate buffer containing 50 mM mPEG-succinimidyl valerate (SVA) (Laysan Bio Inc.) and 1 mM biotin-PEG-SVA (Laysan Bio Inc.) to the coverslips, placing them onto the slides, and incubating at RT in the dark overnight. Finally, flow chambers were disassembled, rinsed with milliQ water, blow-dried, reassembled, and sealed using double-sided tape (3M) and 5-minute epoxy (Devcon).

In order to lyse the cells, they were first centrifuged at 800g for 5 minutes and resuspended in lysis buffer (10 mM Tris pH 7.4, 1 mM EDTA (Fisher Scientific), 1 mM benzamidine (Santa Cruz Biotechnology), 10 µg/mL leupeptin (Sigma-Aldrich), and 1% n-dodecyl-b-d-maltoside (DDM) (Chem-Impex International, Inc.)). This was incubated in a rotator at 4°C for 30 minutes and then centrifuged at 13000g for 30 minutes. The supernatant was collected as the lysate to be used in SiM-Pull experiments.

SiM-Pull experiments were performed in the chambers described above by first rinsing the flow chamber with wash buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl), incubating 0.5 mL of 0.2 mg/mL streptavidin (Promega) in wash buffer for 10 minutes, then flushed with 1 mL wash buffer followed by 1 mL binding buffer (10 mM Tris-HCl, 50 mM NaCl, 0.1 mg/mL bovine serum albumin). The surface was prepared for pull-down by incubating 0.5 mL of 10nM biotinylated anti-mouse (Sigma-Aldrich) in binding buffer for 20 minutes, flushing with 2 mL of binding buffer, incubating 0.5 mL of 10 nM mouse anti-mNeonGreen (Chromotek) for 20 minutes, flushing with 2 mL of binding buffer once more, and finally incubating 1

mL of binding buffer containing the lysate of approximately 5000 cells for 20 minutes. Excess lysate was cleared from the chamber prior to imaging by a final flush with 1 mL binding buffer.

Single molecule imaging

Imaging was performed using the optical system detailed in Appendix 1. A 100x 1.49 NA objective was used for imaging and photobleaching (Olympus UApoN 100x/1.49). The band pass emission filters in the detection pathway were ET525/41 for mNG and ET605/52m for mSc. Cameras were aligned prior to two color imaging using a glass bottom dish with 0.1µm TetraSpeck beads (Thermo Fisher Scientific) adsorbed to the surface at a concentration of about ~300 beads per FOV. Imaging was performed at room temperature.

Image Analysis

Stepwise photobleaching data was analyzed using Progressive Image Filtering (discussed in depth in Appendix 2) with settings that produced results closest to the approximately 5000 manually scored traces [27].

Fluorescence colocalization was performed using a custom Matlab script (also discussed in depth in Appendix 2) with the colocalization threshold set to 1.5 pixels, a value used in previous SiM-Pull experiments in literature [26].

Substrate Preparation

To minimize glass autofluorescence and maximize cell attachment, plain glass-bottom dishes were cleaned and coated with fibronectin. Dishes were etched with 1 M KOH for 20 minutes, followed by a water rinse. For fibronectin coating, dishes were incubated in 4% (3-Mercaptopropyl) trimethoxysilane (Sigma-Aldrich)

in ethanol for 30 minutes, rinsed with ethanol, incubated with (N-γ-maleimidobutyryl-oxysuccinimide ester) crosslinker (4 mM in ethanol, Thermo Scientific), rinsed with ethanol and dried thoroughly in a sterile biosafety cabinet. Dishes were then incubated with 5 µg/mL fibronectin for 2 hours at room temperature, followed by overnight at 4°C, then rinsed with PBS and stored in PBS at 4°C until use (up to several weeks).

Fluorescence Lifetime Imaging Microscopy and Förster Resonance Energy Transfer

Fluorescence lifetimes and lifetime image were measured using time-correlated single photon counting fluorescence measurements (TCSPC), which were carried out using ~120 fs pulses at 900 nm delivered at an 80 MHz repetition rate from a Spectra-Physics Mai-Tai Ti:S laser equipped with DeepSee dispersion compensation. The Ti:S laser was coupled to a Zeiss 880 laser scanning microscope which was used to locate and focus on the cells. A 40x 1.1 NA C-APO water immersion Zeiss objective was used for FLIM imaging. Two-photon generated fluorescence was separated from the excitation using a 670 nm long pass dichroic filter, which directed the emission to a GaAsP photomultiplier tube after passing through a 530/30 band-pass filter (Chroma Technology Corp,). The laser power was attenuated using a near infrared (NIR) Acousto Optic Modulator (AOM) to keep the photon detection rate to less than 0.2% of the repetition rate to avoid photon pile-up. Fluorescence lifetime images were collected using a high-resolution TCSPC FLIM module (SPC-830, Becker & Hickl GmbH) and FLIM images fit to a bi-exponential decay function modified to take the 12.5 ns pulse

repetition timing into account. FLIM data was fit using the SPCImage software package (Becker & Hickl GmbH).

Two-color Fluorescence Cross-Correlation Spectroscopy

FCCS data of mNeonGreen and mScarlet labeled mGluRs was acquired using a Zeiss LSM880 confocal microscope using 488 and 561 nm excitation delivered through a 40/1.2 C-APO water immersion objective lens. The pinhole was set for 1.0 AU and two emission channels (500-550 nm and 580-620 nm) were used to collect the photon counts. Auto and cross-correlation data was acquired and fit using the FCS mode of the Zeiss Zen software.

REFERENCES

1. Pin, J.P. and R. Duvoisin, *The metabotropic glutamate receptors: structure and functions.* Neuropharmacology, 1995. **34**(1): p. 1-26.
2. Niswender, C.M. and P.J. Conn, *Metabotropic glutamate receptors: physiology, pharmacology, and disease.* Annu Rev Pharmacol Toxicol, 2010. **50**: p. 295-322.
3. Doumazane, E., et al., *A new approach to analyze cell surface protein complexes reveals specific heterodimeric metabotropic glutamate receptors.* FASEB J, 2011. **25**(1): p. 66-77.
4. Romano, C., W.L. Yang, and K.L. O'Malley, *Metabotropic glutamate receptor 5 is a disulfide-linked dimer.* J Biol Chem, 1996. **271**(45): p. 28612-6.
5. Kammermeier, P.J., *Functional and pharmacological characteristics of metabotropic glutamate receptors 2/4 heterodimers.* Mol Pharmacol, 2012. **82**(3): p. 438-47.
6. Choi, D.W., *Glutamate neurotoxicity and diseases of the nervous system.* Neuron, 1988. **1**(8): p. 623-34.
7. Aiba, A., et al., *Deficient cerebellar long-term depression and impaired motor learning in mGluR1 mutant mice.* Cell, 1994. **79**(2): p. 377-88.
8. Gil-Sanz, C., et al., *Involvement of the mGluR1 receptor in hippocampal synaptic plasticity and associative learning in behaving mice.* Cereb Cortex, 2008. **18**(7): p. 1653-63.

9. Pekhletski, R., et al., *Impaired cerebellar synaptic plasticity and motor performance in mice lacking the mGluR4 subtype of metabotropic glutamate receptor.* J Neurosci, 1996. **16**(20): p. 6364-73.

10. Linden, A.M., et al., *Increased anxiety-related behavior in mice deficient for metabotropic glutamate 8 (mGlu8) receptor.* Neuropharmacology, 2002. **43**(2): p. 251-9.

11. Duvoisin, R.M., et al., *Increased measures of anxiety and weight gain in mice lacking the group III metabotropic glutamate receptor mGluR8.* Eur J Neurosci, 2005. **22**(2): p. 425-36.

12. Brody, S.A., F. Conquet, and M.A. Geyer, *Disruption of prepulse inhibition in mice lacking mGluR1.* Eur J Neurosci, 2003. **18**(12): p. 3361-6.

13. Brody, S.A., et al., *Assessment of a prepulse inhibition deficit in a mutant mouse lacking mGlu5 receptors.* Mol Psychiatry, 2004. **9**(1): p. 35-41.

14. Dolen, G. and M.F. Bear, *Role for metabotropic glutamate receptor 5 (mGluR5) in the pathogenesis of fragile X syndrome.* J Physiol, 2008. **586**(6): p. 1503-8.

15. Dolen, G., et al., *Correction of fragile X syndrome in mice.* Neuron, 2007. **56**(6): p. 955-62.

16. Beqollari, D. and P.J. Kammermeier, *Venus fly trap domain of mGluR1 functions as a dominant negative against group I mGluR signaling.* J Neurophysiol, 2010. **104**(1): p. 439-48.

17. Kunishima, N., et al., *Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor.* Nature, 2000. **407**(6807): p. 971-7.

18. Robbins, M.J., et al., *Characterization of the dimerization of metabotropic glutamate receptors using an N-terminal truncation of mGluR1alpha.* J Neurochem, 1999. **72**(6): p. 2539-47.

19. Selkirk, J.V., et al., *Characterization of an N-terminal secreted domain of the type-1 human metabotropic glutamate receptor produced by a mammalian cell line.* J Neurochem, 2002. **80**(2): p. 346-53.

20. Levitz, J., et al., *Mechanism of Assembly and Cooperativity of Homomeric and Heteromeric Metabotropic Glutamate Receptors.* Neuron, 2016. **92**(1): p. 143-159.

21. Yin, S., et al., *Selective actions of novel allosteric modulators reveal functional heteromers of metabotropic glutamate receptors in the CNS.* J Neurosci, 2014. **34**(1): p. 79-94.

22. Sevastyanova, T.N. and P.J. Kammermeier, *Cooperative signaling between homodimers of metabotropic glutamate receptors 1 and 5.* Mol Pharmacol, 2014. **86**(5): p. 492-504.

23. Lundstrom, L., et al., *Structural determinants of allosteric antagonism at metabotropic glutamate receptor 2: mechanistic studies with new potent negative allosteric modulators.* Br J Pharmacol, 2011. **164**(2b): p. 521-37.

24. Shaner, N.C., et al., *A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum.* Nat Methods, 2013. **10**(5): p. 407-9.

25. Suh, Y.H., K. Chang, and K.W. Roche, *Metabotropic glutamate receptor trafficking.* Mol Cell Neurosci, 2018. **91**: p. 10-24.

26. Jain, A., et al., *Probing cellular protein complexes using single-molecule pull-down.* Nature, 2011. **473**(7348): p. 484-8.

27. McGuire, H., et al., *Automating single subunit counting of membrane proteins in mammalian cells.* J Biol Chem, 2012. **287**(43): p. 35912-21.

CHAPTER 5
CONCLUSIONS AND FUTURE OUTLOOK

Single molecule imaging methods have fundamentally changed the way we view biology, both literally and figuratively. The identity of, force generated by, and interaction between proteins is a picture that is slowly but steadily being revealed, one experiment at a time. The vast array of tools that fall under the "single molecule imaging method" umbrella are able to measure complex stoichiometry, spatial arrangement of individual molecules, interactions between proteins and DNA, intramolecular conformational changes, intermolecular events, changes in chemical environment and molecular mass, most of which can be observed in real time and some can even be measured or observed *in vivo*. Furthermore, super resolution methods, which have largely been omitted from discussion in this dissertation, continue their march to resolutions even further below the diffraction limit, using more colors simultaneously and at timescales that are increasingly relevant to biological processes.

DNA-protein interactions are critical to the function and proliferation of single and multicellular organisms. Several decades of research, using both single molecule methods and bulk biochemical techniques, have yielded a wealth of information regarding transcription factors, DNA repair mechanisms, gene location and identity, and nuclear architecture. However, much about the nucleus, especially the changes in architecture surrounding gene expression, remains a mystery due to the difficulty of observing and measuring the residents without disrupting them in such a tightly packed and active environment. "DNA curtains" provides a platform on which the dynamics of translocases, helicases, mismatch

repair complexes, and the processes underlying homologous recombination and targeted gene editing have been studied. The method of fabricating these flow cells has been cost prohibitive and relatively inaccessible to a large fraction of the scientific community due to the need for electron-beam lithography tools. I have demonstrated in this dissertation that photolithography, which utilizes tools more widely available, can be employed to pattern the DNA curtains flow cells so that this platform is made available to a wider array of researchers. I found that Transcription Activator-Like Effectors (TALEs) are not compatible with the DNA curtains platform, but the Green, Finkelstein, and Redding labs have demonstrated that many proteins are suitable for use with the technique. A recent innovation by the Redding lab titled “chromatin curtains” has enabled the observation of factors interacting with chromatin, a more biologically relevant substrate than naked DNA, at the level of spatial and temporal detail that the “DNA curtains” platform affords [1].

While some single molecule imaging techniques are currently applicable *in vivo*, one of the main drawbacks of the others is the inability to do so. Improving these techniques for use *in vivo* is facilitated by continued innovation in fluorescent protein and organic dye design, detector technology, and methods for diluting or isolating areas so that single molecules can be observed. In this dissertation, I have shown that we developed a method, titled Single Protein Recovery After Dilution (SPReAD), which utilizes viral mediated membrane fusion as a tool to dilute cellular contents allowing techniques such as single molecule Förster resonance energy transfer (smFRET), fluorescence correlation spectroscopy

(FCS), stepwise photobleaching, and fluorescence colocalization to be performed *in vivo*. In order to induce fusion, our technique requires that cells undergo a brief incubation in pH 5.5 buffer. We demonstrated that this does not have a significant impact on intracellular pH, but there are viral proteins which induce membrane fusion without a change in pH that may prove to be viable alternatives [2].

Understanding the stoichiometry of membrane receptors *in vivo* is critical to developing pharmacological agents that influence their behavior. In this dissertation, I have demonstrated that SPReAD can be used to prepare cells for single molecule imaging of metabotropic glutamate receptors *in vivo.* Metabotropic glutamate receptors mediate the neuromodulatory effects of glutamate in the central nervous system and have been linked to numerous pathologies including mental retardation and schizophrenia. We found that mGluRs labeled with an N-terminal mScarlet were unable to form dimers with N-terminal mNeonGreen labeled mGluRs. Since mScarlet is a poor candidate for stepwise photobleaching, we were unable to determine if mScarlet labeled homodimers occur. Further experiments are necessary to elucidate whether it is the location of mScarlet in the fusion protein that is the issue, if mScarlet interacts with mGluRs in a detrimental way regardless of the location, or if this inability to form dimers is due to another reason. A potential concern for heterodimer formation in transduced or transfected cells is the numerous post translational modifications that mGluRs undergo prior to membrane trafficking [3]. If each subunit of the heterodimer is processed in a separate area of the endoplasmic reticulum, then they will not have the chance to interact in order to form their characteristic sulfur bridge.

Single molecule imaging methods continue to evolve and our understanding of biology at the molecular level does so alongside it. Single molecule imaging provides real time observation and measurement in a relatively non-invasive manner and bridges the gap between biochemical understanding of cellular processes and the detailed molecular structure of proteins provided by structural biology. The upshot of this unique combination is that single molecule methods will remain a powerful and flexible technique suitable for tackling the challenges involved in building an understanding of life from individual molecules.

REFERENCES

1. E. Jensen, L., B. Al-Sady, and S. Redding, *Chromatin Curtains: A Single-Molecule Method for Visualizing Histone Marks on Chromatin Substrates In Vitro*. Vol. 114. 2018. 254a.
2. Harrison, S.C., *Viral membrane fusion.* Virology, 2015. **479-480**: p. 498-507.
3. Suh, Y.H., K. Chang, and K.W. Roche, *Metabotropic glutamate receptor trafficking.* Mol Cell Neurosci, 2018. **91**: p. 10-24.

APPENDIX A

DESIGN OF A CUSTOM MICROSCOPE FOR SINGLE MOLECULE IMAGING

The single molecule imaging performed in this dissertation (excluding the fluorescence correlation spectroscopy (FCS), fluorescence recovery after photobleaching (FRAP), and fluorescence lifetime imaging microscopy (FLIM)) was performed using a microscope that I redesigned and assembled. We required a system with several specialized features: 1) equipped with multiple laser lines, 2) capable of rapidly modulating multiple illumination wavelengths, 3) capable of epifluorescent, highly inclined and laminated optical sheet (HILO), and objective total internal reflection fluorescence (TIRF) illumination, 4) possessing the ability to detect and separate two unique emission wavelengths at the low signal levels that accompany single molecule imaging.

To this end, I redesigned and built a custom microscope with two detection pathways capable of circle scanning objective TIRF, SAIM, STORM, PALM, and other single molecule techniques utilizing widefield or TIRF illumination [1]. Illumination is provided by any combination of the six available excitation laser lines corresponding to commonly used fluorophore absorption wavelengths: 405 nm (Laserglow), 488 nm (SpectraPhysics), 532 nm (Laserglow), 561 nm (Crystalaser), 640 nm (Obis), and 730 nm (Star-tech). These are combined coaxially using 5 dichroic mirrors (Chroma). The beam then passes through a collinear acousto-optical tunable filter (AOTF) (AA optoelectronics) capable of rapidly switching between, modulating the intensity of, or combining laser lines depending on the requirements of the imaging experiment.

The AOTF achieves this by generating an acoustic wave on the order of 100 MHz in a TeO_2 crystal causing an oscillating mechanical strain which, due to photoelasticity, results in a spatial pattern of alternating indices of refraction [2]. These alternating regions act as a diffraction grating, the geometry of which is determined by the frequency and amplitude of the acoustic wave. By aligning the optical path to the first order diffraction maxima, the experimentalist is able to control the intensity of each incident wavelength that is transmitted to the rest of the optical system by altering the properties of the acoustic wave generated by the AOTF.

An ultrafast Uniblitz shutter (Vincent electronics) enables rapid mechanical shuttering of all sample illumination. The beam is directed onto a set of orthogonally oriented galvanometers (6215H, Cambridge Technologies) driven at ~1 KHz by a Micromax 678 dual axis galvanometer driver (Cambridge Technologies) controlled by a scan controller card designed and built in our laboratory.

The user interfaces with this via Visual C++ software designed and written in our laboratory. This allows the experimentalist to change the pattern scanned by the laser, the frequency of rotation, and the x-y location of the beam on the back of the objective. For the ideal usage of circle scanning objective TIRF, the beam should complete several rotations within each imaging frame, so that any optical abnormalities in the illumination are averaged over. Common values used for this are 1KHz oscillation with 100ms frames, averaging 100 complete beam oscillations.

A beam telescope positioned after the galvanometers produces a collimated beam of the appropriate diameter entering a final lens which focuses the beam on the back aperture of the objective lens, creating a collimated beam exiting the objective lens that is necessary for TIRF and SAIM. The optical elements between the galvanometers and the microscope body are secured in a 60 mm cage system (Thorlabs) to ensure that the lenses are aligned and parallel to one another and provide a rigid structure that makes the components and alignment more resistant to vibrations and impacts.

The imaging itself is done on an IX-81 (Olympus) microscope body with a Flat-Top XYZ automated stage (MS-2000, ASI) and various sample holder inserts, some of which were machined in lab. A filter cube is present in the body of the microscope to ensure the separation of the excitation illumination and collected emission into their appropriate optical paths. The system is currently equipped with three filter configurations: 405/488/532/641 nm, 405/488/561/641 nm, and a single 561 nm long pass filter.

The emission pathway consists of a TuCam (Andor) to split individual wavelengths, determined by the filter cube mounted in the TuCam, and directed into two iXion 897 (one of which is the Ultra model) EMCCD cameras (Andor). The current emission filter configurations are optimized for 1) GFP and RFP (525/41 + 580 nm dichroic mirror + 605/52), 2) GFP and Alexa-647 (525/40 + 580 nm dichroic mirror + 679/451), and 3) RFP and Alexa-647 (605/52 + 640 nm dichroic mirror + 705/72) (All emission filters from Semrock, IDEX Health and Science). Imaging and camera control are controlled by Andor Solis (Andor).

The entire optical setup is housed on an ST series EQ damping optical table with S-2000 high performance laminar flow isolators (Newport) to minimize the effect of outside vibrations.

REFERENCES

1. Colville, M.J., et al., *High-speed device synchronization in optical microscopy with an open-source hardware control platform.* bioRxiv, 2019: p. 533349.
2. Debye, P. and F.W. Sears, *On the Scattering of Light by Supersonic Waves.* Proceedings of the National Academy of Sciences, 1932. **18**(6): p. 409-414.

APPENDIX B
IMAGE ANALYSIS METHODS AND ALGORITHMS

Accurate image analysis is critical to optical microscopy because it bridges the divide between the images gathered during an experiment and the quantifiable data from which statistically significant conclusions can be drawn. The ability to accurately and reliably automate this analysis is of growing importance since the continued development of high throughput techniques increases the amount of data that can be gathered in a short time. Herein I have outlined some common techniques used in image analysis and reported the effect of varying the values of parameters in Progressive Image Filtering, a program designed to automate the analysis of stepwise photobleaching data. It was determined that step-to-noise ratio, step dwell time, and the tolerance of step height variation were the parameters which had the most significant effect on the results. These results were then compared to those gathered by two researchers independently scoring traces manually and to another program designed by our lab entitled ImageC. The automated values and the manually scored traces did not match.

INTRODUCTION

One of the advantages of fluorescence microscopy is direct observation of the phenomenon being studied rather than analyzing information inferred from a gel shift or measured reaction product. However, quantifying the data is a nontrivial process with many potential roads to travel before reaching the end. The responsible researcher must be diligent in selecting a method that most accurately quantifies the data while avoiding any personal bias. There are a number of techniques for spatial and temporal image analysis, such as

background reduction, noise filtering, edge detection, particle tracking, fluorophore bleaching and blinking detection. Some enable automated evaluation which eliminates human error and can greatly speed up the rate at which data is processed.

I will discuss the analysis of data generated by two experiments, stepwise photobleaching and fluorescence colocalization, as they both are related to determining the stoichiometry of protein complexes [1, 2]. Before I can discuss the algorithms, I must address some of the filters they utilize, and because an in depth examination of edge detection methods alone would be a lengthy document, I will focus on filters which directly apply: the Gaussian spatial filter, the Laplacian second order derivative operator for edge detection, and the Chung-Kennedy filter for reducing noise in the intensity vs time traces [3, 4].

METHODS

Laplacian of a Gaussian filter

The Laplacian of a Gaussian (LoG) filter is the combination of a Gaussian smoothing filter to reduce noise and a Laplacian filter which is a spatial second order derivative operation for edge detection such as those of a fluorescent puncta or structure in an image [5]. The combination is convenient because of its associative property. The Gaussian and Laplacian filters can be convolved with one another to generate a kernel. This kernel can then be applied to the image, eliminating the need to run each filtering operation on the image individually [3]. Gaussian smoothing is a common way to reduce noise in image analysis. The value of each pixel is converted into a weighted average of itself and neighboring

pixels by convolving an N x N kernel approximating the values of a 2-D Gaussian with the image. The 2-D Gaussian distribution has the form

$$G(x,y) = \frac{1}{2\pi\sigma^2}\exp(-\frac{(x-\mu_x)^2 + \left(y-\mu_y\right)^2}{2\sigma^2}) \quad (1)$$

with μ_x and μ_y as the centroid location and σ as the standard deviation. A 3x3 Gaussian filter kernel with σ = 1 is

$$G = \frac{1}{16}\begin{bmatrix}1 & 2 & 1\\ 2 & 4 & 2\\ 1 & 2 & 1\end{bmatrix} \quad (2)$$

where the 1/16 is a scaling factor so the result of the convolution equals unity. One can see that the value of the pixel to which this is applied will be the average of the surrounding pixels, eliminating high frequency background noise. The Laplacian filter is one of a number of second order derivative methods for edge detections which work by finding changes in sign in the second derivative. The Laplacian operator is given by

$$L(x,y) = \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} \quad (3)$$

The equation describing the LoG filter arises when the Laplacian (3) is applied to a Gaussian function (1),

$$LoG(x,y) = \frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right]\exp(-\frac{(x-\mu_x)^2 + \left(y-\mu_y\right)^2}{2\sigma^2}) \quad (4)$$

A 5x5 LoG kernel with σ=1 is

$$LoG = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & -16 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (5)$$

where the sum of the elements must be zero to result in a zero valued convolution. Edges are detected in the image when the value of the convolution crosses zero. In the case of single molecule image analysis, the desired edges are fluorescent puncta.

Chung-Kennedy filter

The Chung-Kennedy filter is a temporal filter which serves to reduce the noise in time traces while preserving meaningful steps in the data such as a photobleaching event in our case[4]. This is a nonlinear digital filter which looks at points both forwards and backwards of the current time value in order to avoid averaging over a significant feature contained within the noise. First consider a recorded trace y(t) such that

$$y(t) = x(t) + n(t) \quad (6)$$

with x(t) as the desired signal and n(t) as some noise component. A forward and backward predictor for x(t) at time t = T are given by (7) and (8) respectively

$$\hat{x}^f(T) = \frac{1}{3}(y(T-1) + y(T-2) + y(T-3)) \quad (7)$$

And

$$\hat{x}^b(T) = \frac{1}{3}(y(T+1) + y(T+2) + y(T+3)) \quad (8)$$

In a system with no underlying features and uniform noise,

$$\hat{x}^b(T) \cong \hat{x}^f(T) \cong \hat{x}(T) \quad (9)$$

but in an interesting system, these must be considered separately due to the potential presence of a significant feature, so

$$\hat{x}(T) = f(T)\hat{x}^f(T) + b(T)\hat{x}^b(T) \quad 10)$$

where f(T) and b(T) are weights inversely proportional to their χ^2. Thus, if a bleaching step occurs at time t > T then the backwards predictor will have a large variance and the forward predictor should be used to calculate the value of x(T) so that the feature is not included in the calculation of x(T). This is done using a number of differently sized bins as smaller bin sizes will be sensitive to high frequency changes such as a photobleaching event, while larger bin sizes will be sensitive to broad features such as a slow exponential signal decay.

Stepwise photobleaching data is notoriously difficult to analyze in an automated manner due to the non-uniformity of step height, the stochastic nature of bleaching and blinking, small step heights relative to noise, and exponential background signal intensity decay. Even manually scoring the number of steps in each trace can be inconsistent between evaluators (Fig 1a). Ignoring all but the most obvious of bleach steps would avoid ambiguity, but that has an inherent bias toward smaller step numbers as complexes with a larger number of fluorophores tend to exhibit smaller step sizes, have a larger chance of two blinking events happening in close temporal proximity, and hence, the issue of blinking becomes even more complex. I will now introduce two automated analysis software packages: Progressive Image Filtering (PIF) and ImageC, the latter of which was developed in our laboratory [6].

Progressive Image Filtering

The first automated method of counting photobleaching steps I will discuss is Progressive Image Filtering (PIF) which utilizes both of the above concepts in the pursuit of finding the true number of photobleaching steps in each trace[6]. First, a user defined number of frames of the movie are averaged in order to improve the signal to noise and develop a background level. This image is then subjected to LoG filtering in order to reduce noise and locate spots based on a user defined threshold for intensity and either a symmetric or asymmetric Gaussian filter of width (σ). After all of the spots matching these criteria have been identified, the average intensity of the remaining pixels is measured for each frame in order to fit the exponential background decay. This exponential is then subtracted from each trace prior to photobleach step counting.

To crudely summarize the algorithm used by PIF, it analyzes each trace by continuing to combine intensity levels until there are no changes larger than a user defined threshold. I would advise the reader to look at the publication introducing PIF for a thorough explanation, but I will briefly outline the process here, as it relates to several of the parameters I investigated [6].

The algorithm sets discrete intensity levels by finding a point that is either above the value at t = 0, or below it by some threshold, then averaging that point with the initial t = 0 value, and setting the whole interval to that value. This is repeated until the end of the time trace, yielding some number of flat steps of varying length in time. If the length of time spent at a particular level is shorter than user defined threshold, the step is ignored and it is combined into the

previous level. This process is then repeated for larger threshold values until the user defined threshold is reached. A final test of the image was performed to ensure that each step height was larger than some multiple of the standard deviation of the trace, and the image was discarded if this test was failed.

When dealing with such noisy and visually ambiguous data sets, having a number of parameters at your disposal can help with designing a fit that best matches the gold standard of evaluation: the number of steps measured manually by a human. However, the drawback to having so many dimensions of parameter space is that the data can be either accidentally, or intentionally, biased to support or disprove a hypothesis. Furthermore, examining every combination in parameter space to find the “ideal” is laborious, especially if it needs to be re-evaluated frequently for different data sets. I examined the axes of parameter space by affixing all but one parameter while varying the remaining one and recording the bleach step histogram results for the same 10 images.

The parameters I explored were: symmetric vs asymmetric Gaussian filter geometry (Figure B.1a), the full width at half maximum (FWHM) of the Gaussian that was fit to each spot (Figure B.1b), the number of iterations the algorithm will perform before reaching the full minimum step amplitude value (Figure B.1c), the tolerance of variation between the magnitude of different steps (Figure B.1d), the minimum time the trace was required to dwell at a level to be considered a step (Figure B.1e), and the minimum ratio of step amplitude vs standard deviation of noise (Figure B1f). Changing the values for (a), (b), and (c) had a minor effect on

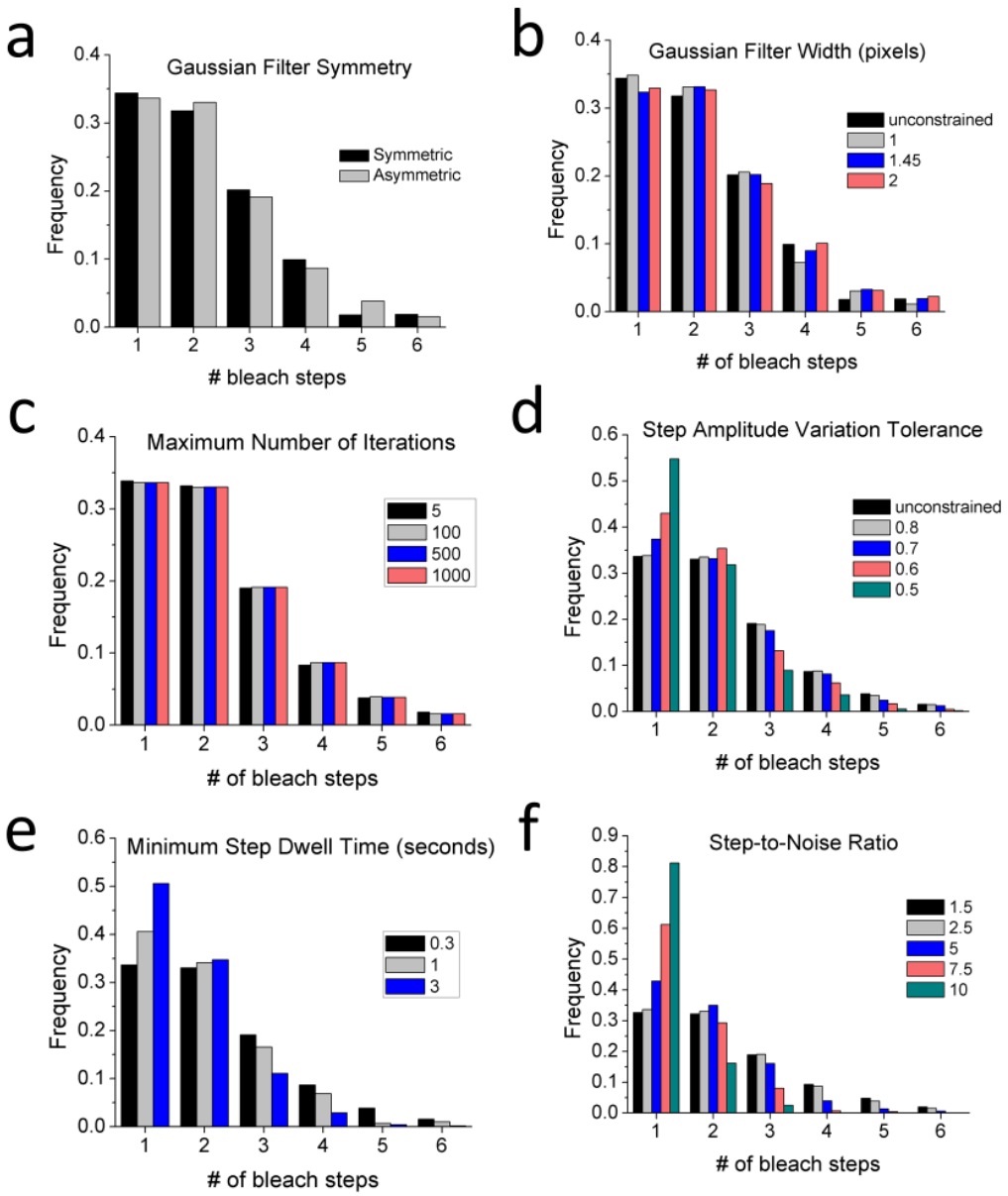


Figure B.1 Determining the Impact of PIF Parameters
(a) The geometry of the Gaussian filter used to find and accept or reject spots. (b) The FWHM of the Gaussian used to find and accept or reject spots, when unconstrained the algorithm would select the FWHM that best fit each spot. (c) The maximum number of iterations the algorithm will run before reading out the step count. (d) The tolerance for variation in step amplitudes, when unconstrained this is no longer a step rejection parameter. (e) Minimum time the trace is required to stay at a given intensity level in order for it to be considered a valid level. (f) Ratio of step height to standard deviation of noise.

the resulting step distribution, while (d), (e), and (f) had a significant effect on the resulting step distribution. This result was expected as the latter three parameters correspond to the three features the human eye catches when manually scoring traces: step height uniformity, time occupying a given intensity level, and the signal-to-noise ratio.

ImageC

I will now discuss the algorithms used by ImageC before I return to a comparison of the two software packages. There is not a publication detailing ImageC, so I will discuss it more thoroughly than PIF. ImageC locates spots by scanning an N x N region across the image, where N is a parameter set by the user. If a pixel in the region is above the user defined threshold, it is fitted to an N x N Gaussian filter. The minimum goodness of fit (assuming a perfect Gaussian) is defined by the user, and any spots that fall below the limit are rejected. Intensity vs time traces for each spot which met the criteria are next loaded into the time trace analysis portion of the software (Figure B.2a). Many images contain some level of background fluorescence that follows an exponential decay which is quantified by measuring the intensity over time of all “non-spot” pixels (Figure B.2b). This exponential is then subtracted from the time trace to simplify step counting. Prior to searching for steps, a linear fit of the time trace is analyzed to eliminate those which have a small change in intensity. The maximum slope (more negative slopes correspond to traces which have a greater decrease in intensity) is defined by the user. In order to recognize steps,

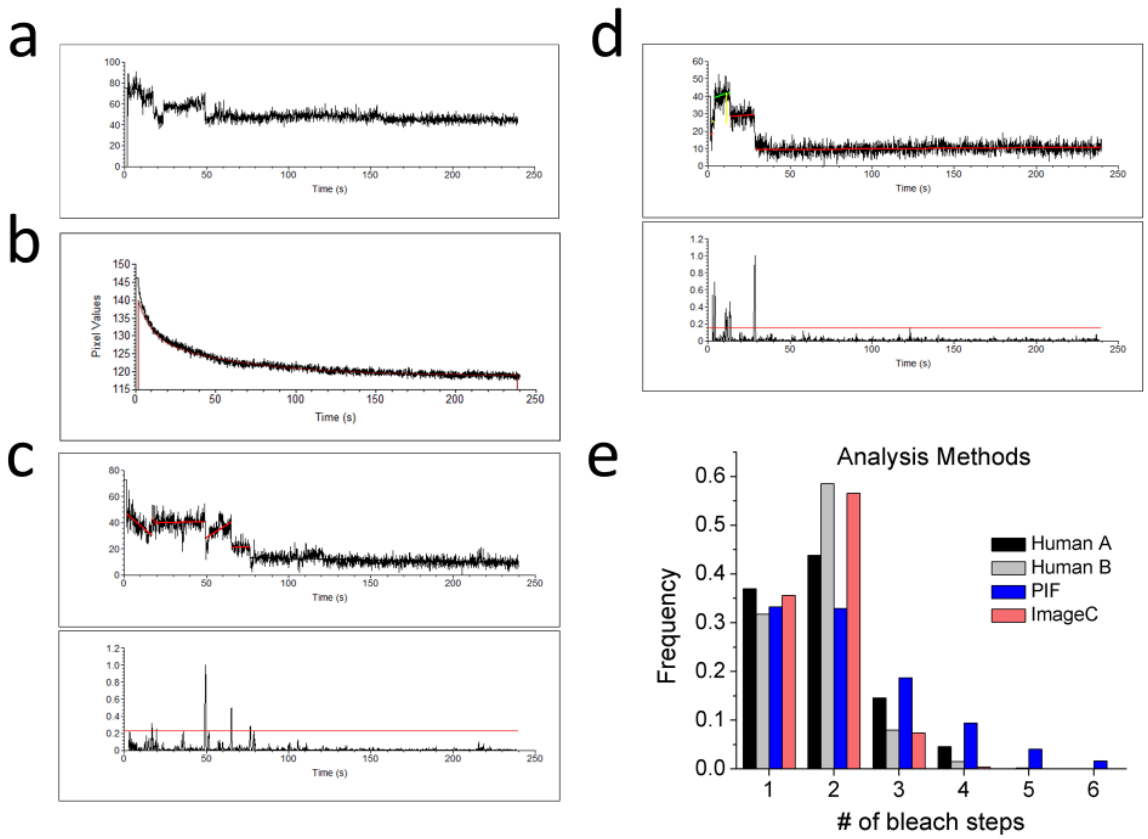


Figure B.2 ImageC and Stepwise Photobleaching Data Analysis Comparison
(a) Example of a stepwise photobleaching trace demonstrating how ambiguous some steps can be. Time trace plotted in ImageC. (b) Fit of background exponential decay using ImageC. (c) Example of the trace and incorrect "one bleach step" fit by ImageC (above) and the plot of $kdata^2$ (below). (d) Trace showing two clear bleach steps (three levels) accurately scored in ImageC (above) along with $kdata^2$ (below). (e) Bleach step histogram of two manual analyses and two automated analyses of the same set of 10 images containing ~4000 points in total.

a kernel with a user defined width, of the form

$$K = \begin{cases} 1 \; for \; k_i < \dfrac{M}{2} \\ -1 \; for \; k_i > \dfrac{M}{2} \end{cases} \tag{11}$$

is convoluted with the time trace, and the time derivative of the trace is approximated by *kdata(t)* which is given by

$$kdata(t) = \sum_{i=1}^{M} k_i (c_i - \frac{1}{M} \sum_{j=1}^{M} c_i) \tag{12}$$

with k_i is the value of the kernel at *i* and c_i is the value of the time trace. This value is then squared to accentuate the peaks and convert all of the elements to positive numbers. Peaks, referred to as “Kernel jogs”, in this derivative trace correspond to bleaching events, blinking events, or areas with large variations in background. The maximum number of these “kernel jogs” is set by the user, corresponding to the maximum number of steps that are expected to be observed. The threshold is automatically set at a level that finds the maximum number of jogs. These “jogs” should contain all of the bleach steps plus other events. The segments between these “jogs” are then considered intensity levels, and they are fit to a simple linear approximation. The maximum slope allowed for this is set by the user and exceeding this slope disqualifies that level. The minimum time at which the intensity remains at that level is also set by the user. The deviation of the trace about this line is recorded and used as criteria to evaluate the surrounding “jogs” as possible steps. The “jog” is disqualified if

$$S_i(end) - fsd < S_{i+1}(start) + fsd \tag{13}$$

or

$$S_i(end) + fsd > S_{i+1}(start) - fsd \quad (14)$$

with S_i as the *i*th segment, *S(start)* as the first point in the linear fit of *S*, *S(end)* as the last point in the linear fit of *S*, and *fsd* is the fractional standard deviation about the linear fit (Figure B.2d). The number of bleach steps in the trace is set as the number of "jogs" remaining after this final evaluation. Both software packages score some traces incorrectly. Sometimes these traces are ambiguous to the human eye, and at other times they would be obvious. Figure B.2c and d show an improperly scored and properly scored trace respectively.

ImageC and PIF both use a similar algorithm to find spots, evaluate the time spent at a level, and calculate deviation of the trace about a linear fit of the level in order to determine a step. The major difference lies in the method of detection of a possible step. PIF defines a potential bleach event as a change in intensity larger than some threshold, while ImageC defines a potential bleach event as a point with a large time derivative.

Comparing and Evaluating Algorithms and Manual Analysis

A data set containing 10 images with ~400 points per image was analyzed manually by myself (Human A), Nitya Deshmukh (Human B) a member of the lab, PIF, and ImageC (Figure B.2e). The subjective nature of time trace scoring can be observed by the difference in the bleach step histogram between Humans A and B. ImageC accurately replicated the results of Human B, and PIF did so with Human A. This is an interesting finding and was done so with minimal adjustments of the parameters in ImageC. A thorough exploration of parameter

space in ImageC, as was done for PIF, will be conducted to see how each parameter affects the results.

Fluorescence Colocalization

A second method of measuring the stoichiometry is fluorescence colocalization. As discussed in Chapter 1 of this dissertation, the three metrics one can use to quantify colocalization are Manders' and Pearson's coefficients, and spot based methods measuring distance between centroids. I'll discuss the latter in further depth in this chapter as it is the method I chose for colocalization analysis.

Object based colocalization is a very intuitive measure of the relationship between two fluorescent molecules. The centroid of each fluorescent puncta is found by fitting its intensity profile to a Gaussian and then, using mNeonGreen and mScarlet as an example, the distances between a green centroid and every red centroid is measured and then this distance measurement is repeated for every green centroid. These distances are compared to a user defined threshold, often on the order of 1-2 pixels (~100-250nm), and if two spots are within that radius, they are considered to be colocalized. The fraction of colocalization is the number of colocalized puncta divided by the smaller of the two numbers of fluorophores.

These algorithms are calibrated by using synthetic plots with a predetermined number of overlaps and measuring the colocalization of that image when permutated, such as by rotation or reflection relative to an axis. Random distributions of the number of points found in experimental data can

provide a baseline number from which a positive or negative deviation can be measured and labeled proteins can be inferred to be part of a complex, strictly exclusive, or any form of interaction between.

Furthermore, the optical elements, most notably the cameras, used to gather the data need to be aligned, and the use of fiducial markers can further adjust for misalignment. Specially crafted calibration slides and slides containing fluorescent beads are the commonly used standards for both physically aligning the cameras and measuring if any shift or rotation is necessary in the analysis code (Figure B.3a). Images of randomly distributed beads are useful as negative controls to ensure the analysis code isn't returning false positives (Figure B.3b)

My algorithm, in particular, finds all of the puncta that exceed a unique user defined intensity threshold for each color and each image. These puncta are then fit to a Gaussian and if the fit is good and the FWHM is approximately 1.7 pixels (~180 nm corresponds to the Rayleigh limit for a point source emitting at ~550 nm), the centroid of each is recorded. To adjust for any slight camera misalignment, one of the images is rotated in 0.004 rad steps to -0.04 and 0.04 rad, and translated from -2 to +2 pixels in steps of 0.5 pixels in both x and y at each value of rotation (Figure B.3c). The best x-y shift for each rotation step is recorded and yields information regarding the direction in which the cameras are misaligned. Camera alignment is performed by imaging a bead slide with both cameras and making adjustments until the puncta are coregistered between the cameras. The colocalization should be nearly 100%, so if no value of rotation and translation within the aforementioned ranges exceed 85%, the images from that

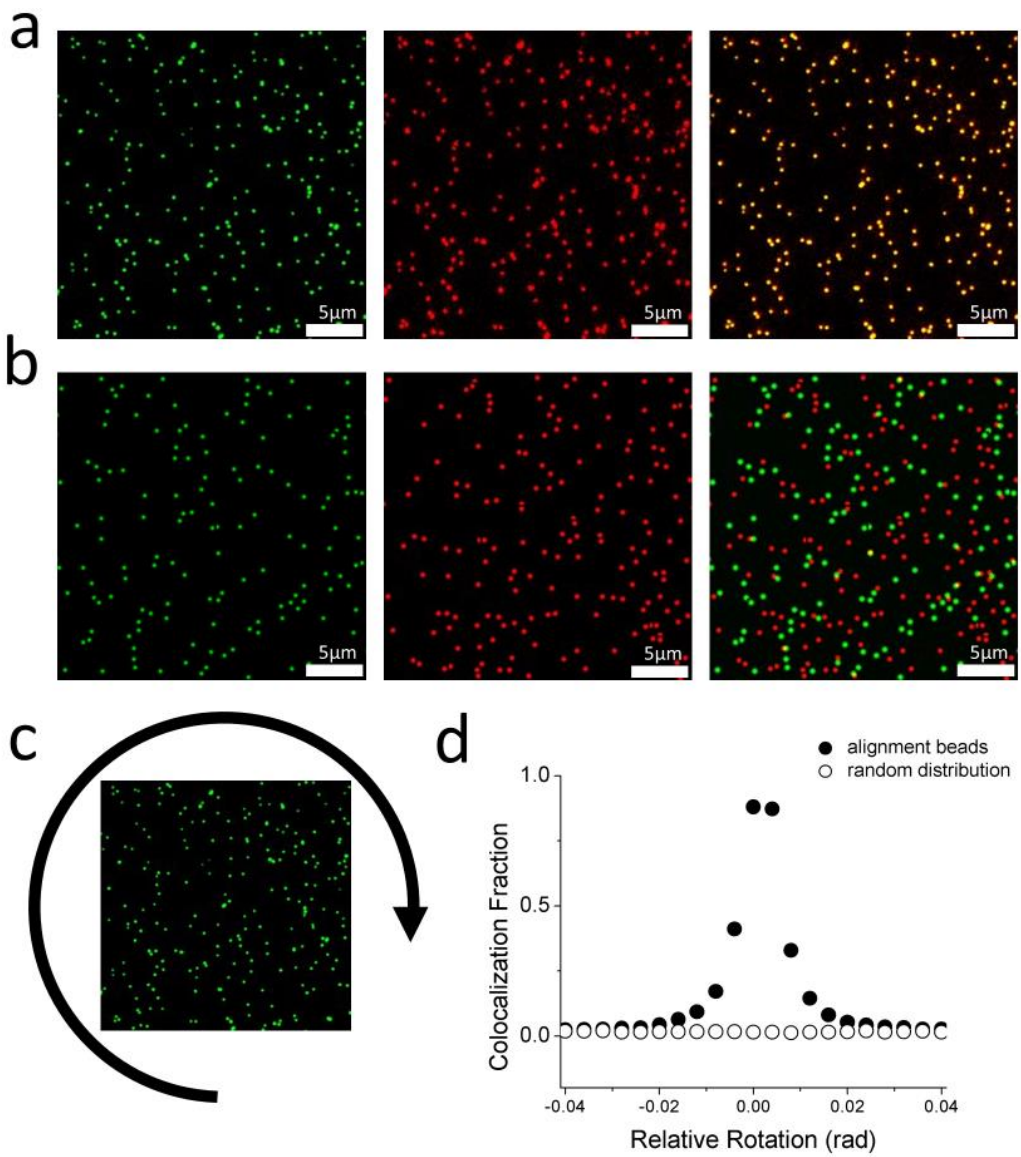


Figure B.3 Colocalization Algorithm Explanation and Testing
(a) Colocalization of a bead slide containing 0.1µm TetraSpeck beads using the Matlab script described in this Appendix. Green emission (left panel) and red emission (middle panel) overlap with an 85% colocalization (right panel). (b) Images of a random distribution of 600 Gaussian spots generated by a Matlab script. The two channels are pseudo-colored red and green for the sake of consistency with the rest of the figure (left and middle panel respectively). The colocalization fraction was found to be 4%, as expected for a field of random points. (c) Diagram demonstrating the function of the Matlab script which rotates and shifts one channel to adjust for any camera misalignment. (d) Plot showing the fraction of colocalization as the image is rotated for both TetraSpeck beads and the random distribution of points.

day are considered unusable. Furthermore, the fraction of colocalization should drop off sharply as the image is rotated away from the maximum value in each direction affirming that the colocalization is a significant result (Figure B.3d).

CONCLUSION

Accurate analysis of both stepwise photobleaching and fluorescence colocalization data is required to draw accurate conclusions based on the data gathered. Stepwise photobleaching, in particular, would benefit from accurate and precise automated analysis software since the data sets are large and manually scoring ambiguous traces can vary over time. The consistent nature of automated scoring produces data with more confidence, as it is, at the very least, consistent in its inaccuracy. I have outlined several algorithms used by two software packages for automated analysis of stepwise photobleaching in this appendix and explored the parameter space of PIF. This will be repeated for ImageC in an effort to find the optimal set of parameters to match manually scored traces or highlight ways that the program could be improved.

REFERENCES

1. Ulbrich, M.H. and E.Y. Isacoff, *Subunit counting in membrane-bound proteins.* Nat Methods, 2007. **4**(4): p. 319-21.
2. Lagache, T., et al., *Statistical analysis of molecule colocalization in bioimaging.* Cytometry A, 2015. **87**(6): p. 568-79.
3. Juneja, M. and P. Sandhu, *Performance Evaluation of Edge Detection Techniques for Images in Spatial Domain*. Vol. 1. 2009. 614-621.
4. Chung, S.H. and R.A. Kennedy, *Forward-backward non-linear filtering technique for extracting small biological signals from noise.* Journal of Neuroscience Methods, 1991. **40**(1): p. 71-86.
5. Bhardwaj, S. and A. Mittal, *A Survey on Various Edge Detector Techniques.* Procedia Technology, 2012. **4**: p. 220-226.
6. McGuire, H., et al., *Automating single subunit counting of membrane proteins in mammalian cells.* J Biol Chem, 2012. **287**(43): p. 35912-21.